LIBRARY
OF THE
UNIVERSITY of ILLINOIS.

ON THE CONVERGENCE OF CONTINUED FRACTIONS WITH COMPLEX ELEMENTS*

BY

E. B. VAN VLECK

Up to the present time few theorems of a general character for the convergence of continued fractions with complex elements have been obtained, and these few are of very recent date. In the first section of this paper such theorems upon the subject as are known to the writer are brought together for the purpose of indicating the present state of our knowledge, and the scope of the paper is also explained. Some new criteria for convergence are then deduced in the succeeding sections. The results obtained are summed up in theorems 1–10, which may be read independently of the rest of the paper. The demonstration of these theorems is based upon certain equations, Nos. 3–8, 11, and 12, which seem to be new and of a fundamental character.

I. *Summary of previous criteria for convergence.*

§ 1. Two classes of continued fractions

$$\frac{\mu_1}{\lambda_1}+\frac{\mu_2}{\lambda_2}+\frac{\mu_3}{\lambda_3}+\cdots$$

with real elements are commonly treated in mathematical writings. In the first class all the elements are positive; in the second the numerators of the partial quotients are negative and the denominators positive.

Complete criteria have been obtained for the convergence and divergence of continued fractions of the first kind. If, namely, the fraction is thrown into the form

$$\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}+\cdots \qquad (a_n \geqq 0),$$

it will converge when $\sum a_n$ is divergent; on the other hand, if $\sum a_n$ is convergent, the even convergents have one limit and the odd convergents have another. No criteria of equal generality have been obtained for the second class of con-

* Presented to the Society February 23, 1901. Received for publication April 15, 1901.

50037

tinued fractions. The principal theorem, however, establishes the convergence of the fraction when $\lambda_n \geqq 1 + |\mu_n|$.*

§ 2. The convergence-theorems for continued fractions with imaginary elements may be regarded as generalizations of these theorems for real continued fractions. In 1898 PRINGSHEIM† showed that the condition $|\lambda_n| \geqq 1 + |\mu_n|$ was sufficient to ensure the convergence not only of continued fractions with real elements but also of those containing complex elements. If, furthermore, the continued fraction has the customary normal form in which $\mu_n = 1$, this condition may be replaced by the less restricted one:

$$\frac{1}{|\lambda_{2n-1}|} + \frac{1}{|\lambda_{2n}|} \leqq 1.$$

In this modified form PRINGSHEIM's criterion is applicable to the continued fractions of HURWITZ,‡ in which the partial denominators are complex integers.

The only extension of the convergence-theorem for the first class of real continued fractions is contained in a remarkable and extended memoir by STIELTJES,§ published in 1894. STIELTJES considers the fraction

$$\frac{1}{a_1 z} + \frac{1}{a_2} + \frac{1}{a_3 z} + \frac{1}{a_4} + \cdots \qquad (a_n > 0, \quad n = 1, 2, 3, \cdots),$$

and shows that when $\sum a_n$ is divergent, it converges over the entire plane of z with the exception of the negative half of the real axis. On the other hand, when $\sum a_n$ is convergent, the even and the odd convergents separately converge, the limits of their numerators and denominators being holomorphic functions of *genre* 0 whose roots all lie upon the negative half of the axis.

From a function-theoretical standpoint this result is of great importance. It also calls attention anew to the fact that algebraic continued fractions commonly have a more extended region of convergence than the corresponding infinite series, and hence have, at least prospectively, a wide field of usefulness. The convergence-proofs of STIELTJES are most admirable, both by themselves and in relation to the theory of functions. When, however, they are considered solely with respect to the general theory of continued fractions, they are not entirely satisfactory. Not only are his proofs based upon the properties of the conver-

* For reference to the possibilities when this condition is not fulfilled, see PRINGSHEIM's report in the *Encyklopädie der mathematischen Wissenschaften.*

† Sitzungsberichte der mathematisch-physikalischen Classe der Münchener Akademie, vol. 28, p. 295.

‡ Acta Mathematica, vol. 11.

§ Annales de la Faculté des Sciences de Toulouse, vol. 8, with continuation in vol. 9.

gents as functions of z,* but the form of his continued fraction is also very restricted, inasmuch as the ratio of the real to the imaginary component is the same in all the denominators of the alternate partial quotients.

These considerations have led to the present paper. It comprises such results as have suggested themselves in seeking to extend the convergence-theorems for the first class of real continued fractions to complex continued fractions.

To complete the historical review, mention should be made of the theorems of PRINGSHEIM and STIELTJES relating to continued fractions in which the denominators of the partial quotients are equal to unity. PRINGSHEIM† proves that such a fraction will converge if

$$|\mu_2| < \tfrac{1}{2}, \quad |\mu_{2n+2}| + |\mu_{2n+1}| \leqq \tfrac{1}{2} \qquad (n = 1, 2, \cdots).$$

STIELTJES‡ treats the fraction

$$\frac{b_1 z}{1} + \frac{b_2 z}{1} + \frac{b_3 z}{1} + \cdots \qquad (b_n > 0),$$

and shows that it represents a meromorphic function when b_n approaches zero as its limit, and only then.

Various other theorems pertaining to continued fractions are reducible to the foregoing by transformation of the fraction.

II. *Fundamental equations.*

§ 3. Let

$$\frac{\mu_1}{\lambda_1} + \frac{\mu_2}{\lambda_2} + \frac{\mu_3}{\lambda_3} + \cdots$$

be any continued fraction whose elements are complex numbers. Denote by p_n and q_n the numerator and denominator of the n-th convergent, by M_n the modulus of q_n. If then we put

$$\lambda_n = a_n + i\beta_n, \quad \mu_n = \gamma_n + i\delta_n,$$

$$p_n = r'_n + is'_n, \quad q_n = r_n + is_n,$$

the familiar laws for the formation of the successive convergents,

$$(1) \qquad p_n = \lambda_n p_{n-1} + \mu_n p_{n-2}, \quad q_n = \lambda_n q_{n-1} + \mu_n q_{n-2},$$

*See, for example, the remarkable preliminary theorem which he establishes in order to prove the convergence of the continued fraction when $\sum a_n$ is divergent and the real part of z is negative. Loc. cit., vol. 8, p. J56.

†Sitzungsberichte der Münchener Akademie, vol. 28.

‡Annales de Toulouse, vol. 9, p. A, 42–47.

may be replaced by

$$r_n = a_n r_{n-1} + \gamma_n r_{n-2} - \beta_n s_{n-1} - \delta_n s_{n-2},$$
$$s_n = a_n s_{n-1} + \gamma_n s_{n-2} + \beta_n r_{n-1} + \delta_n r_{n-2}, \tag{2}$$

with two similar equations for r_n' and s_n'.

Squaring and adding equations (2) we obtain next the following fundamental relation between the moduli of three consecutive denominators:

$$M_n^2 = (a_n^2 + \beta_n^2) M_{n-1}^2 + (\gamma_n^2 + \delta_n^2) M_{n-2}^2 + 2(a_n\gamma_n + \beta_n\delta_n) C_{n-1} + 2(\beta_n\gamma_n - a_n\delta_n) D_{n-1}, \tag{3}$$

in which

$$C_{n-1} = r_{n-1}r_{n-2} + s_{n-1}s_{n-2}, \tag{4}$$

$$D_{n-1} = r_{n-1}s_{n-2} - s_{n-1}r_{n-2}. \tag{5}$$

The form of the last two equations may be modified advantageously by introducing in place of r_{n-1} and s_{n-1} the expressions which are obtained by the substitution of $n-1$ for n in (2). The equations then become

$$C_{n-1} = a_{n-1}M_{n-2}^2 + \gamma_{n-1}C_{n-2} - \delta_{n-1}D_{n-2}, \tag{6}$$

$$D_{n-1} = -\beta_{n-1}M_{n-2}^2 - \gamma_{n-1}D_{n-2} - \delta_{n-1}C_{n-2}. \tag{7}$$

In this form they may be used with (3) to compute M_{n-1}, C_{n-1}, D_{n-1} from the corresponding quantities of next lower subscript. If also we take

$$r_0 = 1, \quad r_0' = s_0 = s_0' = 0,$$

we shall have

$$M_0^2 = 1, \quad C_1 = r_1 = a_1 M_0^2, \quad D_1 = -s_1 = -\beta_1 M_0^2.$$

Hence C_{n-1}, D_{n-1}, M_{n-1}^2 are each homogeneous linear functions of the squares of the moduli M_i with lower subscripts (ultimately of M_0^2 and M_1^2), in which the coefficients are entire functions of a_i, β_i, γ_i, $\delta_i (i = 1, 2, \cdots, n-1)$.

From (4) and (5) we also get the important relation:

$$C_{n-1}^2 + D_{n-1}^2 = M_{n-1}^2 M_{n-2}^2. \tag{8}$$

§ 4. To investigate the convergence of a continued fraction recourse is sometimes had to the equation:*

$$\frac{p_{n+m}}{q_{n+m}} = \frac{p_n}{q_n} + (-1)^n \mu_1\mu_2\cdots\mu_{n+1}\left(\frac{1}{q_n q_{n+1}} - \frac{\mu_{n+2}}{q_{n+1}q_{n+2}} + \frac{\mu_{n+2}\mu_{n+3}}{q_{n+2}q_{n+3}} + \cdots + (-1)^{m-1}\frac{\mu_{n+2}\mu_{n+3}\cdots\mu_{n+m}}{q_{n+m-1}q_{n+m}}\right), \tag{9}$$

* HEINE'S *Kugelfunctionen*, vol. 1, p. 263.

in which m is allowed to increase indefinitely. The convergence of a continued fraction is thus made to depend upon that of an infinite series. As, however, this form of the series is not well adapted to the consideration of cases of alternating convergence, it will be better to group by twos the terms in the parenthesis of (9). With the aid of the relation

$$\frac{1}{q_n q_{n+1}} - \frac{\mu_{n+2}}{q_{n+1} q_{n+2}} = \frac{\lambda_{n+2}}{q_n q_{n+2}},$$

the series may then be thrown into the form:

$$\frac{p_n}{q_n} + (-1)^n \mu_1 \mu_2 \cdots \mu_{n+1} \left(\frac{\lambda_{n+2}}{q_n q_{n+2}} + \frac{\lambda_{n+4} \mu_{n+2} \mu_{n+3}}{q_{n+2} q_{n+4}} + \frac{\lambda_{n+6} \mu_{n+2} \mu_{n+3} \cdots \mu_{n+5}}{q_{n+4} q_{n+6}} + \cdots \right). \tag{10}$$

Two different cases arise according as n is even or odd. If the series converges in both cases and the continued product of the μ_i is convergent and greater than zero, *the necessary and sufficient condition for the coincidence of the two limiting values of the series is that* $M_{n-1} M_{n-2}$ *shall increase indefinitely with* n. The continued fraction then converges in the ordinary sense of the term. On the other hand, if $M_{n-1} M_{n-2}$ has a finite limit, the convergence is of an alternating character.

The series (10) has the advantage over (9) that it is frequently absolutely convergent when the latter is only conditionally convergent.

Either series loses its significance if q_i vanishes in one or more of its terms. Should this, however, happen only a finite number of times, the continued fraction may still converge, for the objectionable terms can be removed from the series by increasing n sufficiently.

In one other particular the convergence of a continued fraction requires special comment. The removal of the first r partial quotients μ_i / λ_i may change a convergent continued fraction into a divergent one, or *vica versa*. Adopting a term proposed by PRINGSHEIM,* we shall say that the convergence is *unconditional* when it is not affected by the removal of any finite number of partial quotients, beginning with the first.

§ 5. Besides discussing the convergence of the fraction, we shall in certain cases determine the signs of its real and imaginary parts. For this purpose let p_n / q_n be thrown into the form:

$$\frac{r_n r_n' + s_n s_n' + i(r_n s_n' - r_n' s_n)}{M_n^2}.$$

* For a further discussion of unconditional convergence consult the memoir of PRINGSHEIM previously cited.

If we then place

$$K_n = r_n r'_n + s_n s'_n,$$
$$L_n = r_n s'_n - r'_n s_n,$$

and combine these equations with (2) and the two analogous equations for r'_n and s'_n, we obtain

$$K_n = (a_n^2 + \beta_n^2) K_{n-1} + (\gamma_n^2 + \delta_n^2) K_{n-2} + 2(a_n\gamma_n + \beta_n\delta_n) k_{\alpha\gamma}^{(n-1)} + 2(\beta_n\gamma_n - a_n\delta_n) k_{\beta\gamma}^{(n-1)}, \tag{11}$$

$$L_n = (a_n^2 + \beta_n^2) L_{n-1} + (\gamma_n^2 + \delta_n^2) L_{n-2} + 2(a_n\gamma_n + \beta_n\delta_n) l_{\alpha\gamma}^{(n-1)} + 2(\beta_n\gamma_n - a_n\delta_n) l_{\beta\gamma}^{(n-1)}, \tag{12}$$

in which

$$k_{\alpha\gamma}^{(n-1)} = \tfrac{1}{2}(r_{n-1}r'_{n-2} + r'_{n-1}r_{n-2} + s_{n-1}s'_{n-2} + s'_{n-1}s_{n-2}),$$
$$k_{\beta\gamma}^{(n-1)} = \tfrac{1}{2}(r_{n-1}s'_{n-2} + r'_{n-1}s_{n-2} - s_{n-1}r'_{n-2} - s'_{n-1}r_{n-2}),$$
$$l_{\alpha\gamma}^{(n-1)} = \tfrac{1}{2}(r_{n-1}s'_{n-2} - r'_{n-1}s_{n-2} - s_{n-1}r'_{n-2} + s'_{n-1}r_{n-2}),$$
$$l_{\beta\gamma}^{(n-1)} = \tfrac{1}{2}(-r_{n-1}r'_{n-2} + r'_{n-1}r_{n-2} - s_{n-1}s'_{n-2} + s'_{n-1}s_{n-2}).$$

The last set of equations may be simplified in the same manner as (4) and (5). They then become

$$k_{\alpha\gamma}^{(n-1)} = a_{n-1}K_{n-2} + \gamma_{n-1}k_{\alpha\gamma}^{(n-2)} - \delta_{n-1}k_{\beta\gamma}^{(n-2)},$$
$$k_{\beta\gamma}^{(n-1)} = -\beta_{n-1}K_{n-2} - \gamma_{n-1}k_{\beta\gamma}^{(n-2)} - \delta_{n-1}k_{\alpha\gamma}^{(n-2)},$$
$$l_{\alpha\gamma}^{(n-1)} = a_{n-1}L_{n-2} + \gamma_{n-1}l_{\alpha\gamma}^{(n-2)} - \delta_{n-1}l_{\beta\gamma}^{(n-2)},$$
$$l_{\beta\gamma}^{(n-1)} = -\beta_{n-1}L_{n-2} - \gamma_{n-1}l_{\beta\gamma}^{(n-2)} - \delta_{n-1}l_{\alpha\gamma}^{(n-2)},$$

and may be used in conjunction with (11) and (12) to find K_n and L_n for successive values of n.

It will, of course, be observed that the above equations are similar in their construction to equations 3–7.

III. *On the convergence and properties of a certain class of continued fractions.*

§ 6. Our attention in this section will be confined entirely to continued fractions which have been reduced to the normal form for which $\gamma_n = 1$, $\delta_n = 0$. Equations (3), (6), and (7) then take the simpler form:

$$M_n^2 = (a_n^2 + \beta_n^2) M_{n-1}^2 + M_{n-2}^2 + 2a_n C_{n-1} + 2\beta_n D_{n-1}, \tag{13}$$

$$C_{n-1} = a_{n-1}M_{n-2}^2 + C_{n-2}, \tag{14}$$

$$D_{n-1} = -\beta_{n-1}M_{n-2}^2 - D_{n-2}, \tag{15}$$

and the series (10) simultaneously becomes

$$\frac{p_n}{q_n} + (-1)^n \left(\frac{a_{n+2} + i\beta_{n+2}}{q_n q_{n+2}} + \frac{a_{n+4} + i\beta_{n+4}}{q_{n+2} q_{n+4}} + \cdots \right). \tag{16}$$

We shall first determine what signs for a_n and β_n are most favorable to the absolute convergence of (16), supposing their absolute values to be given. Since

$$C_{n-1} = \sum_{i=1}^{n-1} a_i M_{i-1}^2, \quad D_{n-1} = -\sum_{i=1}^{n-1} (-1)^{n-i+1} \beta_i M_{i-1}^2, \tag{17}$$

C_{n-1} will increase in absolute value with increasing n if the a_n have a common sign, and D_{n-1} will similarly increase if the β_n are alternately of opposite sign (each zero value of β_n being included and reckoned with an appropriate sign). Moreover $a_n C_{n-1}$ and $\beta_n D_{n-1}$ will then never be negative. These conditions are clearly as favorable as possible for the increase of $M_n = |q_n|$ and hence also for the absolute convergence of (16).

§ 7. In considering the convergence of the continued fraction two cases are to be distinguished according as $\sum |\lambda_n|$—that is, $\sum |a_n + i\beta_n|$—if convergent or divergent.

Consider first the former case and suppose either that the a_n have a common sign or that the β_n alternate in sign. It is immaterial which supposition is made because a_n and $-\beta_n$ enter in like manner into the fundamental equations. In either case we have from (1):

$$M_m \leqq |\lambda_m| M_{m-1} + M_{m-2} \leqq (1 + |\lambda_m|) M',$$

in which M' is used to represent the larger of the two moduli M_{m-1}, M_{m-2}. This shows that

$$M_n < M' \prod_{i=m}^{n} (1 + |\lambda_i|) \qquad (n > m).$$

But by a well-known theorem $\prod (1 + |\lambda_i|)$, $(i = 1, 2, \cdots)$, converges with $\sum |\lambda_i|$. There must accordingly be an upper limit to the value of M_n. On the other hand, by virtue of the hypothesis concerning a_n or β_n, either $|C_{n-1}|$ or $|D_{n-1}|$ will increase with n, or at least not decrease. It follows then from (8) that $M_{n-1} M_{n-2}$ has a lower limit. This can only be if M_n, which has an upper limit, has also a lower limit.

Turning next to the series (16) we see now that it must be absolutely convergent, since the numerators of its terms form an absolutely convergent series while the moduli of the denominators have a lower limit. As also $M_{n-1} M_{n-2}$ does not indefinitely increase, two distinct values for the series will be obtained according as n is odd or even. We have therefore the following result:

THEOREM I. *In any continued fraction*

$$\frac{1}{a_1+i\beta_1}+\frac{1}{a_2+i\beta_2}+\frac{1}{a_3+i\beta_3}+\cdots \tag{18}$$

the even and the odd convergents will both converge and their limits will be distinct if (1) $\sum|a_n+i\beta_n|$ *is convergent and* (2) *either the a-constituents have a common sign or the β-constituents alternate in sign.**

§ 8. In § 7 it was shown, and without making use of condition (1), that $M_{n-1}M_{n-2}$ had a lower limit. An immediate inference from this is that M_n cannot vanish for any finite value of n. A like conclusion obviously holds for the numerator of the nth convergent, inasmuch as the equations for the formation of the successive numerators are similar to those for the denominators of the convergents. We have thus a very simple proof of the following theorem:

THEOREM 2. *Neither the numerator nor the denominator of any convergent of the continued fraction* (18) *will vanish if the a-constituents have a common sign or if the β-constituents alternate in sign.*

§ 9. We pass now to the consideration of the case in which $\sum|a_n+i\beta_n|$ diverges. Let it be first supposed that both a_n and β_n are restricted in sign in the manner indicated in the preceding theorem. Every term in (13) will then be positive or zero, and M_n must be equal to or greater than M_{n-2} for every value of n. Accordingly M_n has a lower limit. Furthermore, by our initial hypothesis either $\sum|a_n|$ or $\sum|\beta_n|$ must diverge, or both. It follows therefore from (17) that at least one of the two quantities, C_{n-1} and D_{n-1} increases indefinitely with n. But by (8) $M_{n-1}M_{n-2}$ must increase in the same manner. If, therefore, the series (16) converges, it must have the same limit whether n be odd or even.

To prove that (16) does in fact converge, we shall establish separately the convergence of the two series:

$$\sum_{n=1}^{\infty}\frac{|a_n|}{M_{n-2}M_n},\quad \sum_{n=1}^{\infty}\frac{|\beta_n|}{M_{n-2}M_n}.$$

Consider first the former. If in computing M_n by formula (13) for successive values of n we should equate each β_n to zero, smaller values for M_n would be obtained, which we will denote by M'_n. Obviously these are successively connected by the relation:

$$M'_n=|a_n|\,M'_{n-1}+M'_{n-2},$$

or by its equivalent

$$\frac{1}{M'_{n-2}M'_{n-1}}=\frac{|a_n|}{M'_{n-2}M'_n}+\frac{1}{M'_{n-1}M'_n}.$$

* Zero-constituents are admissible in this and the following theorems, being taken as having the signs which belong to the places in which they are found.

If in the latter we set in turn $n = m, m+1$ — and add we get,

$$\frac{1}{M'_{m-2}M'_{m-1}} \geqq \sum_{n=m}^{\infty} \frac{|a_n|}{M'_{n-2}M'_n}.$$

This proves that

$$\sum \frac{|a_n|}{M'_{n-2}M'_n}$$

converges, and hence also $\sum |a_n| M_{n-2}^{-1} M_n^{-1}$. Similar considerations apply to prove the convergence of $\sum |\beta_n| M_{n-2}^{-1} M_n^{-1}$.

We conclude, therefore, that the continued fraction has a definite limit, and the reasoning evidently holds if the first n partial quotients are omitted. This result may be restated in

THEOREM 3. *If the a-constituents in* (18) *all have the same sign and the β-constituents alternate in sign, the continued fraction will converge unconditionally when* $\sum |a_n + i\beta_n|$ *is divergent.*

§ 10. The signs of the real and imaginary parts of the convergents can be quickly determined. Equations 13–15 hold when M_i^2, C_i, D_i are replaced by $K_i, k_{\alpha\gamma}^{(i)}, k_{\beta\gamma}^{(i)}$ or by $L_i, l_{\alpha\gamma}^{(i)}, l_{\beta\gamma}^{(i)}$. The initial values of these quantities are as follows:

$$K_0 = L_0 = 0, \quad K_1 = a_1, \quad L_1 = -\beta_1,$$

$$k_{\alpha\gamma}^{(1)} = l_{\beta\gamma}^{(1)} = \tfrac{1}{2}, \quad k_{\beta\gamma}^{(1)} = l_{\alpha\gamma}^{(1)} = 0.$$

An easy mathematical induction then shows that all the terms in (13) after either of the above substitutions have a common sign. From this we get at once

THEOREM 4. *If the a-constituents have a common sign and the β-constituents alternate in sign, the real part of any convergent will have the same sign as* a_1 *and the imaginary part the same sign as* $-\beta_1$.

§ 11. Returning now to the continued fraction of theorem 3, we proceed to consider the effect of the admission of a finite number of a-constituents or of β-constituents which fail to satisfy the conditions there imposed. It is immaterial whether the exceptional elements occur among the a_n or the β_n; let them occur among the former.

When n is increased, there must be a value—call it m—subsequent to which a_n and C_{n-1} have each an invariable sign. If they have the same sign, all the steps of the proof of theorem 3 will apply after the m-th convergent. This will necessarily be the case when $\sum |a_n| M_{n-1}^2 (n = 1, 2, \cdots)$ is divergent.

There remains for consideration only the special case in which $\sum |a_n| M_{n-1}^2$ converges and a_n and C_{n-1} have opposite signs when n exceeds m. Equation (13) then gives

$$M_n^2 \geqq a_n^2 M_{n-1}^2 - 2|a_n C_{n-1}| + M_{n-2}^2,$$

whence by combination with (8) we obtain

$$M_n^2 \geqq M_{n-2}^2 \left(1 - \frac{|a_n C_{n-1}| M_{n-1}^2}{C_{n-1}^2 + D_{n-1}^2}\right)^2. \tag{19}$$

Now since $\sum |a_n| M_{n-1}^2$ converges, $|C_{n-1}|$ has an upper limit. On the other hand, $|D_{n-1}|$, which never diminishes, will have a finite lower limit after some fixed value of n, unless it is equal to zero for every value of n. This exceptional case will not occur if there is an infinite number of values of n for which $|\beta_n| > 0$, as we will suppose, neither can it occur if there is a single such value greater than m, inasmuch as $M_{m+i} > 0$. An upper limit must therefore exist for $|C_{n-1}|/(C_{n-1}^2 + D_{n-1}^2)$. Since also $|a_n| M_{n-1}^2$ diminishes indefinitely with increasing n, the second term of the binomial in the last inequality may be made as small as desired. Let m be taken large enough to ensure that it shall always be less than unity when $n > m$. We then have

$$M_{m+2n} > M_m \prod_i \left(1 - \frac{|a_i C_{i-1}| M_{i-1}^2}{C_{i-1}^2 + D_{i-1}^2}\right) \quad (i = m+2, m+4, \cdots, m+2n).$$

If in the last expression n is indefinitely increased, the product $\prod$ must be convergent since $\sum |a_i M_{i-1}^2|$ is convergent and $|C_{i-1}|/(C_{i-1}^2 + D_{i-1}^2)$ has an upper limit. Moreover, by a well-known theorem the limiting value of $\prod$ must be greater than zero. Hence M_{m+2n} has a lower limit, and the same is true of M_n since m may be taken to be either odd or even.

The convergence of $\sum |a_n|$ may next be quickly argued. This follows from the fact that the terms in the convergent series $\sum a_n M_{n-1}^2$ have ultimately a common sign while M_n has a lower limit. Now by our original hypothesis $\sum |a_n + i\beta_n|$ was a divergent series. This necessitates that $\sum |\beta_n|$ should likewise diverge. But when this diverges, $|D_{n-1}|$ increases indefinitely and with it simultaneously $M_{n-1} M_{n-2}$.

To demonstrate the convergence of the continued fraction, it remains only to prove that (16) converges. This can be quickly done. For, on the one hand, $\sum |a_n| M_{n-2}^{-1} M_n^{-1}$ obviously converges. On the other hand, if each a_n is replaced by zero, $\sum |\beta_n| M_{n-2}^{-1} M_n$ will converge, as we saw in § 9. But from (19) it is clear that the presence of the a-terms in (13), when it does not actually increase the value of M_n, will cut it down in a ratio not exceeding

$$1 : \left(1 - \frac{|a_n C_{n-1}| M_{n-1}^2}{C_{n-1}^2 + D_{n-1}^2}\right).$$

It has also been shown that the product of all such ratios for which n exceeds m is convergent. The restoration of the a-terms will therefore not reduce the ~~If in the latter we set in turn $n = m, m+1, \cdots$, and then add, we get~~

value of $M_n (n > m)$ by more than a certain fixed part of itself. Since this in no wise affects the convergence of $\sum |\beta_n| M_{n-2}^{-1} M_n^{-1}$, the series must still converge. This completes the proof.

The conclusion reached can be summed up as follows:

THEOREM 5. *The continued fraction of theorem* 3 *will still converge if a finite number of the* a_n *or of the* β_n *are admitted which fail to satisfy the conditions there imposed, provided that after the last irregularity there is at least one value of* n *for which* β_n, *respectively* a_n *is not equal to zero.*

§ 12. The criterion of theorem 5 differs from most or all criteria hitherto given for the convergence of continued fractions in that it admits a finite number of irregularities whose positions in the continued fraction are entirely arbitrary. This raises the question whether an infinite number of exceptional values of a_n might not have been admitted into the continued fraction, or, in other words, whether the condition imposed upon β_n was not in itself sufficient to ensure the convergence of the continued fraction, as in the case when $\sum |a_n + i\beta_n|$ converged. That this, however, is in no wise true, can be shown by constructing divergent fractions in which all the β-constitutents fulfill the conditions imposed in the theorem.

Consider, for example, a continued fraction in which β_n is equal to zero for an infinite number of values of n but not for all values. When $\beta_n = 0$, we have

$$M_n^2 = \left(M_{n-2} + \frac{a_n C_{n-1}}{M_{n-2}} \right)^2 + \frac{a_n^2 D_{n-1}^2}{M_{n-2}^2}.$$

Let now a_n be so chosen that the term in the parenthesis shall vanish. This makes $|D_{n-1}^{-1}| = |a_n| M_{n-2}^{-1} M_n^{-1}$. If the values of the β_n which do not vanish fulfill the requirements concerning their signs and are also so chosen that $|D_{n-1}|$ shall have an upper limit, the terms of (16) cannot all decrease indefinitely with increasing n. The series (16) and the continued fraction are in consequence divergent.

§ 13. The restriction upon the signs of a_n in theorem 3 may, however, be dispensed with if we assume that $|\beta_n| / |a_n|$ has a finite lower limit. The proof of the convergence of the continued fraction is then not so simple.

The first step will be to show that $|D_n|$ increases indefinitely with n. Suppose, if possible, that the contrary is the case. It will be recalled that the signs of the β_n were so restricted that $|D_n| = \sum |\beta_n| M_{n-1}^2$. If $\sum |\beta_n| M_{n-1}^2$ converges, $\sum |a_n| M_{n-1}^2$ must likewise converge in consequence of the above assumption. Now it was shown in § 11 that when the latter converged, M_n had a finite lower limit. It follows that $\sum |\beta_n|$ should be convergent. But, on the other hand,

the assumption of a lower limit for $|\beta_n|/|a_n|$ and the hypothesis that $\sum|a_n + i\beta_n|$ diverges, together necessitate the divergence of $\sum|\beta_n|$. This involves a contradiction. We conclude, therefore, that $|D_n|$ must increase indefinitely, as stated. Equation (8) shows that $M_{n-1}M_{n-2}$ also increases without limit.

To establish the convergence of (16) it clearly suffices to demonstrate the convergence of $\sum|\beta_n|M_{n-2}^{-1}M_n^{-1}$. Let k^{-1} be a lower limit to the value of $|\beta_n|/|a_n|$. Equations (17) show at once that $|C_{n-1}| \leqq k|D_{n-1}|$, whence it follows by (8) that

$$|D_{n-1}| > \frac{M_{n-2}M_{n-1}}{k+1}.$$

When a_n and C_{n-1} are of like sign for any value of n, we have from (13),

$$M_n^2 > \beta_n^2 M_{n-1}^2 + 2\beta_n D_{n-1} + M_{n-2}^2.$$

The combination of this inequality with the preceding gives

$$M_n > \frac{|\beta_n|}{k+1} M_{n-1} + M_{n-2}, \tag{20}$$

or

$$\frac{1}{M_{n-2}M_{n-1}} > \frac{|\beta_n|}{(k+1)M_{n-2}M_n} + \frac{1}{M_{n-1}M_n}. \tag{21}$$

The last inequality will also hold when a_n and C_{n-1} have opposite signs if the expression $a_n^2M_{n-1}^2 + 2a_nC_{n-1}$ in (13) is positive. The only possible exceptions therefore occur when a_n differs in sign from C_{n-1} and

$$|a_n| < \frac{2|C_{n-1}|}{M_{n-1}^2}.$$

In this case, if we put

$$a_n = -(1+x_n)\frac{C_{n-1}}{M_{n-1}^2}, \tag{22}$$

the absolute value of x_n will be less than unity. The combination of the last equation with

$$C_n = -a_nM_{n-1}^2 + C_{n-1}$$

then gives

$$C_n = -x_nC_{n-1}.$$

Hence the necessary and sufficient condition that $a_n^2M_{n-1}^2 + 2a_nC_{n-1}$ *shall be negative is that* $|C_n|$ *shall be less than* $|C_{n-1}|$.

Even when this condition is fulfilled, the inequality (21) may yet hold. This will, for example, be the case whenever $|C_{n-1}|$ falls below a certain fixed part of $|D_{n-1}|$. For suppose $C_{n-1} < k'|D_{n-1}|$, where k' denotes the positive number less than unity which satisfies the equation

$$\frac{1-kk'}{1+k'}=\frac{1}{1+k}.$$

Then simultaneously

$$|a_n|<k|\beta_n|, \qquad |D_{n-1}|>\frac{M_{n-2}M_{n-1}}{1+k'}.$$

Since also $kk'<1$, it follows that

$$2(a_nC_{n-1}+\beta_nD_{n-1})>2|\beta_nD_{n-1}|(1-kk')>\frac{2|\beta_n|M_{n-1}M_{n-2}}{1+k}.$$

The combination of the last inequality with (13) shows immediately that (20) and (21) also hold.

When for any value of n the inequality (21) is not valid, we shall substitute the expression (22) for a_n in equation (13). This equation then takes the form:

$$M_n^2=\beta_n^2M_{n-1}^2+2\beta_nD_{n-1}+M_{n-2}^2-\frac{(1-x_n^2)C_{n-1}^2}{M_{n-1}^2}.$$

Since also

$$\frac{C_{n-1}^2}{M_{n-1}^2}=\frac{C_{n-1}^2}{C_{n-1}^2+D_{n-1}^2}M_{n-2}^2,$$

this gives

$$M_n^2>\beta_n^2M_{n-1}^2+\frac{2|\beta_n|}{k+1}M_{n-1}M_{n-2}+y_n^2M_{n-2}^2,$$

in which

$$y_n=\sqrt{1-\frac{(1-x_n^2)C_{n-1}^2}{C_{n-1}^2+D_{n-1}^2}}. \tag{23}$$

It follows then *a fortiori* that

$$\frac{1}{M_{n-2}M_{n-1}}>\frac{|\beta_n|}{(k+1)M_{n-2}M_n}+\frac{y_n}{M_{n-1}M_n}. \tag{24}$$

§ 14. For every value of n there is, then, an inequality of the form (21) or (24). Let these inequalities be arranged in order for successive values of n. When thus arranged they may be divided into sets according to the value of $|D_n|$ in the following manner: Let c be any fixed value greater than unity. Denote also by n_2 the first value of n for which $|D_n|\geqq c|D_1|$, by n_3 the first succeeding value for which $|D_n|\geqq c^2D_1$, etc. Then the first set is to comprise all the inequalities which precede the one in which $n=n_2$; the second set is to begin with the latter inequality and to continue until $n=n_3$, and so on.

Since $|D_n|<M_{n-1}M_n$, the left hand member of the first inequality of the $(r+1)$-th set can not be greater than $1/|D_1|c^r$, and the sum of all such left hand members will therefore be less than a convergent geometrical progression. Consider now any set by itself, and let each inequality contained therein be

multiplied by the product of the y_n which occur in the preceding inequalities of the set. If the inequalities thus modified are added together, each left hand member except the first will be canceled by the last term on the right hand side of the preceding inequality. The sum of the uncanceled left hand members will accordingly be less than the above geometrical progression.

It will be shown immediately that the product of all the y_n which occur in any set has a finite lower limit L independent of the number r which specifies the set. Granting this to be true, if we add together the various sets of inequalities, the sum of all the first terms on the right hand side of the inequalities will be greater than

$$\frac{L}{k+1}\sum_{n=1}^{\infty}\frac{|\beta_n|}{M_{n-2}M_n}.$$

As this sum is less than the above geometric progression, $\sum |\beta_n| M_{n-2}^{-1}M_n^{-1}$ has a finite limit. The same must be true of the series (16). Since also $M_{n-2}M_{n-1}$ increases indefinitely, the limit will be the same whether n is odd or even. The continued fraction will therefore converge.

It remains now only to establish the existence of the lower limit L. Let $m_1, m_2, m_3, \cdots$ denote in order the values of n which specify for any set the inequalities which are of the form (24). From inspection of (23) it is clear that the product of the y_n will have a finite lower limit if

$$\sum \frac{(1-x_n)^2 C_{n-1}^2}{C_{n-1}^2 + D_{n-1}^2} \qquad (n = m_1, m_2, m_3, \cdots) \tag{25}$$

has an upper limit. Since $|D_n|$ never diminishes,

$$C_{n-1}^2 + D_{n-1}^2 > D_{m_1-1}^2$$

for each term of this sum, while in the first term

$$|C_{m_1-1}| < k' |D_{m_1-1}|.$$

We have also previously seen that $C_n = -x_n C_{n-1}$ for each of the above values of n. Hence if the ratio

$$|C_{m_{r+1}-1}| : |C_{m_r}|$$

is denoted by κ_r, the series (25) will term by term be less than

$$(k')^2\{(1-x_{m_1}^2) + (1-x_{m_2}^2)x_{m_1}^2\kappa_1^2 + (1-x_{m_3}^2)x_{m_1}^2 x_{m_2}^2 \kappa_1^2\kappa_2^2 + \cdots\}.$$

The expression here in parenthesis obviously has a value not greater than 1 when each κ_n is placed equal to unity. Hence its original value can not exceed the product of all the κ_n^2 which are greater than unity. Our problem is thus reduced to showing that this product has a common upper limit for the various sets of inequalities. This can be quickly proved. For when $|C_n|$ increases

from $|C_{m_r}|$ to $|C_{m_{r+1}-1}| \equiv \kappa_r |C_{m_r}|$, an increment is given to $|D_n|$ which is at least equal to the increment of $|C_n|$ divided by k. Since also for each of the values of n under consideration $|C_n|$ can not fall below $k'|D_n|$, the sum of the increments given to $|D_n|$ must exceed

$$\frac{k'}{k}|D_{m_1}|\sum(\kappa_r - 1).$$

On the other hand, as soon as this sum reaches $c|D_{m_1}|$, or before, a new set of inequalities must be begun. Hence for each set

$$\sum(\kappa_r - 1) < \frac{kc}{k'}.$$

But the product $\Pi\kappa_r$ has an upper limit simultaneously with $\sum(\kappa_r - 1)$. This completes the proof.

It should be noticed that the values of these upper limits and hence of L depend only upon k and c.

The result which has thus been reached may be recapitulated in the following theorem.

THEOREM 6. *When $\sum|a_n + i\beta_n|$ is divergent, the continued fraction* (18) *will converge if the β-constituents alternate in sign and the ratio $|\beta_n| : |a_n|$ has a finite lower limit,* also if a_n never varies in sign and $|a_n| : |\beta_n|$ has such a limit.*

§ 15. When the second set of conditions of the last theorem are fulfilled and $|a_n| \geqq |\beta_n|$ for every value of n, the signs of the real parts of any convergent and hence of the continued fraction can be determined without difficulty. For consider the equations obtained by substituting K_i, $k_{\alpha\gamma}^{(i)}$, $k_{\beta\gamma}^{(i)}$ for M_i^2, C_i, D_i in (13). It will be found that the only term which varies in sign is $2\beta_n k_{\beta\gamma}^{(n-1)}$, and the absolute value of this term is less than that of $2a_n k_{\alpha\gamma}^{(n-1)}$. The following result will then be evident.

THEOREM 7. *If a_n has a constant sign and $|a_n| \geqq |\beta_n|$ for all values of n, the real part of any convergent and hence of the limit of continued fraction will have the same sign as a_1; if β_n alternates in sign and $|\beta_n| \geqq |a_n|$, the sign of the imaginary part is opposite to that of β_1.*

IV. *On a certain class of algebraic continued fractions.*

§ 16. Application of the results of section III. may now be made to algebraic continued fractions. Consider first any fraction of the form

* "*Untere Grenze*" (WEIERSTRASS); not necessarily a value which the ratio approaches when n increases indefinitely.

$$\frac{1}{d_1 z} + \frac{1}{d_2 z} + \frac{1}{d_3 z} + \cdots$$

in which the d_n are real numbers alternately positive and negative.

Theorem 6 is applicable only when $\sum |d_n|$ is divergent, and shows that the continued fraction converges over the entire plane of z with the exception of the whole or a part of the real axis. Theorem 2 proves that the roots of the numerators and denominators lie only upon the real axis. Above and below the axis the convergents are therefore analytic. That their limit is also an analytic function can be seen from the series (16). For since the individual terms of (16) are analytic, it suffices, by a well-known theorem of WEIERSTRASS,* to prove that the convergence is uniform. Take any circle (R) whose boundary lies wholly above or below the axis, and let $1/k$ denote the smallest value of $|y|/|x|$ for this circle. Then since $|a_n| \leqq k|\beta_n|$ throughout (R), we need only to demonstrate the uniform convergence of $\sum |\beta_n| M_{n-2}^{-1} M_n^{-1}$. If in §§ 13 and 14, k is replaced by k, its maximum value within (R), the reasoning of these paragraphs will be seen to hold for every value of z within the circle or upon its boundary. The number L, which depends only on k and c, is moreover independent of z. If $|D_1|$ denotes the smallest value of $|D_1|$ for the circle (R), the series

$$\frac{L}{k+1} \sum |\beta_n| M_{n-2}^{-1} M_n^{-1}$$

will be less, *term by term*, than a geometrical progression whose first term is $|D_1^{-1}|$ and whose ratio c^{-1} is less than unity. It follows that $\sum |\beta_n| M_{n-2}^{-1} M_n^{-1}$ is uniformly convergent. The continued fraction therefore represents an analytic function.

The position of the zeros of this analytic function can be quickly determined from theorem 2. For when a sequence of rational functions—in the present case the convergents—converge uniformly to an analytic function as their limit, the zeros of this function are the condensation-points of the zeros of the rational fractions.† The analytic function defined by the continued fraction therefore can vanish only for real values of z.

§ 17. When $\sum |d_n|$ is convergent, the even and the odd convergents define distinct functions. Theorem 1 shows that they converge over the entire plane with the possible exception of the real axis. Not only is this true, but in either sequence of convergents the numerators and denominators separately converge, and their limits are holomorphic functions. This may be demonstrated as follows:

* WEIERSTRASS, *Abhandlungen aus der Functionenlehre*, p. 73 ff.; or HARKNESS AND MORLEY, *Analytic Functions*, § 81.

† HURWITZ, Mathematische Annalen, vol. 33, p. 249.

The denominators of the convergents, also the numerators, contain alternately only odd and even powers of z, and in each one the coefficient of z^r, when it does not vanish, consists of certain products of the d_n taken r at a time. By the aid of equation (1) it can be easily shown that the products which figure in the coefficient of z^r in the denominator (numerator) of the nth convergent are included among the products which make up the corresponding coefficient of the second succeeding denominator (numerator). Hence as n is indefinitely increased, the coefficient of any power of z tends to take a definite limiting form. But by virtue of our initial hypothesis $\prod(1 + |d_n|z)$ converges for every finite value of z. It may therefore be expressed as a series $\sum c_r z^r$ which likewise converges over the finite plane, and the coefficient c_r is the limit of the sum of all the partial products of the $|d_n|$ taken r at a time. As c_r is obviously greater than the limit of the coefficient of z^r in either sequence of alternate numerators or denominators, the limits of the latter must be holomorphic functions.

It will be observed that *this conclusion always holds when* $\sum|d_n|$ *is convergent, irrespective of the conditions imposed upon* d_n *in our continued fraction.*

In the special case under discussion the position of the roots of the limiting functions can be inferred from theorem 2 in the same manner as when $\sum|d_n|$ was divergent. Theorem 7 may also be applied for a part of the plane.

Before summing up our results, it will be advantageous to transform the continued fraction so as to make the coefficients of the denominators of the partial quotients positive. We have then the following theorem.

THEOREM 8. *If any algebraic continued fraction has the form*

$$\frac{1}{b_1 z} - \frac{1}{b_2 z} - \frac{1}{b_3 z} - \cdots$$

and every element b_n *is positive, the continued fraction will have the following properties:*

(*a*) *The zeros of the numerators and denominators of its convergents all lie upon the real axis of* z.

(*b*) *If* $\sum|b_n|$ *is divergent, it represents a function which is analytic over the entire plane with the possible exception of the axis and which vanishes only for real values of* z.

(*c*) *If* $\sum|b_n|$ *is convergent, the limits of the numerators and denominators of the even, also of the odd convergents are holomorphic functions whose zeros lie upon the real axis.*

(*d*) *The imaginary part of each convergent or of the limiting function will have a sign opposite to the imaginary part of* z *in the two quarter planes for which* $|y| \geqq |x|$.

The continued fraction of the last theorem can readily be transformed, by neglecting a numerical factor, so as to take either the form

$$\frac{1}{a_1 z'} + \frac{1}{a_2 z'} + \frac{1}{a_3 z'} + \cdots \qquad (a_n > 0),$$

or the form preferred by STIELTJES and cited in § 1. To throw it into the first form one has only to put $z = iz'$. If it is to be reduced to the form of STIELTJES, first put $z = w^{-1}$ and transform so as to remove w from the denominators of the partial quotients. Each numerator after the first becomes equal to $-w^2$. The desired form is then immediately obtained by setting $-w^{-2} = z'$ and clearing the numerators of z'.

§ 18. The considerations by which theorem 8 was established apply with little change to continued fractions having the form

$$\frac{1}{b_1 z + c_1} + \frac{1}{-b_2 z - c_2} + \frac{1}{b_3 z + c_3} + \frac{1}{-b_4 z - c_4} + \cdots$$

in which the coefficients b_n have a common sign. If both $\sum b_n$ and $\sum c_n$ are convergent, $\prod (1 + |b_n z| + |c_n|)$ is also convergent and can be expressed as a power series which converges over the finite plane. As in § 17, the numerators and denominators of the alternate convergents tend to assume a definite limiting form, in which the coefficient of each power of z is the sum of certain products of the b_n and c_n whose moduli are included among the products which make up the coefficient of the corresponding power of z in the series for the above product $\prod$. The limits of the numerators and denominators of the alternate convergents are therefore again holomorphic functions.

If $\sum b_n$ or $\sum |c_n|$ is divergent, some additional restriction must be added. Suppose $|c_n|/|b_n|$ to have an upper limit L. Then at least $\sum b_n$ is divergent. Since

$$\frac{|\beta_n|}{|a_n|} = \frac{|y b_n|}{|x b_n + c_n|} \geqq \frac{|y b_n|}{|x b_n| + L|b_n|},$$

it is evident that $|\beta_n|/|a_n|$ has a finite lower limit for any value of z for which $y > 0$. The conclusions of (b), theorem 8, then follow. Hence

THEOREM 9. *If a continued fraction has the form*

$$\frac{1}{b_1 z + c_1} - \frac{1}{b_2 z + c_2} - \frac{1}{b_3 z + c_3} - \frac{1}{b_4 z + c_4} - \cdots$$

in which the coefficients b_n have a common sign, it will have the following properties:

(a) The roots of the numerators and denominators of its convergents all lie upon the real axis of z.

(b) If $\sum b_n$ and $\sum |c_n|$ are convergent, the limits of the numerators and denominators of the odd, also of the even convergents are holomorphic functions whose zeros lie upon the real axis.

(*c*) *If* $\sum b_n$ *is divergent and* $|c_n|/|b_n|$ *has an upper limit, the continued fraction converges for all values of* z *which are not real, and it represents a function which is analytic over the entire plane with the exception of the whole or a part of the real axis and which vanishes only for real values.*

The reality of the roots of the continued fraction has been previously proved by SYLVESTER,* who also shows that the roots of the denominators of two consecutive convergents alternate with each other. He does not consider the question of the convergence of the continued fraction when the number of elements is infinite.

§ 19. The theorems of III. may also be applied to certain classes of irregular continued fractions. It will suffice here to give as illustrations two applications of this nature.

THEOREM 10. *If the elements of an algebraic continued fraction*

$$\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\frac{1}{\lambda_3}+\cdots$$

are all either positive numbers multiplied by z *or complex numbers* $a_n + i\beta_n$ *in which* $|a_n|/|\beta_n|$ *exceeds some fixed positive number for all values of* n, *the continued fraction will have the following properties:*

(*a*) *None of the roots of the numerators and denominators of its convergents will lie within the half plane in which the real part of* z *is positive, or upon its boundary.*

(*b*) *If* $\sum|\lambda_n|$ *is divergent, the continued fraction represents within this half plane an analytic function which does not vanish at any point in its interior.*

(*c*) *If* $\sum|\lambda_n|$ *is convergent, the limits of the numerators and denominators of the even, also of the odd convergents are holomorphic functions whose roots lie without this half plane.*

The second class of irregular algebraic continued fractions is that in which the even (odd) denominators of the partial quotients are all positive numbers and the odd (even) partial denominators are positive numbers or such numbers multiplied by z. The region of convergence covers the same half plane and the zeros of the numerators and denominators of the convergents and of the limiting functions lie upon the negative half of the axis. The continued fraction of STIELTJES is the special case in which z appears in *every* odd denominator.

WESLEYAN UNIVERSITY, MIDDLETOWN, CONN.

* Philosophical Transactions, vol. 143, part I, p. 497.

GEOMETRY WITHIN A LINEAR SPHERICAL COMPLEX*

BY

PERCEY F. SMITH

Introduction.

The following pages are devoted to the study of a point-sphere correspondence of involutory character, which appears as direct generalization from a certain point of view of the well known point-sphere correspondence arising in a dilatation and the point-point correspondence of spherical inversion. Problems resulting from the establishment by this correspondence of a spherical contact transformation are discussed.

Elsewhere in the Transactions† I have shown the importance to the higher spherical geometry of the involutory contact-transformation here studied, for its relation to the group of all spherical contact transformations is precisely that of spherical inversion to the group of the geometry of reciprocal radii.

The titles of the sections indicate to a great extent their contents.

§1. *Point-sphere correspondence: definition and discussion.*

The *æquatio directrix* in the rectangular point coördinates x, y, z, X, Y, Z,

$$\begin{aligned}\Phi \equiv\ & (X^2+Y^2+Z^2-r_0^2)(x^2+y^2+z^2-r_0^2) \\ & +r_0^2(1-\kappa_0^2)[(X-x)^2+(Y-y)^2+(Z-z)^2]=0\,,\end{aligned} \tag{1}$$

establishes in space R_3 the most general involutory point-sphere correspondence Γ such that the ∞^3 coaxial systems determined each by a point and its associated sphere intersect in a fundamental sphere σ, viz., $x^2+y^2+z^2=r_0^2$. Under (1) the point $P(x, y, z)$ corresponds to the sphere S of center x', y', z' and radius r' where

$$\frac{x'}{x}=\frac{y'}{y}=\frac{z'}{z}=\frac{r_0^2(1-\kappa_0^2)}{x^2+y^2+z^2-r_0^2\kappa_0^2}, \tag{2}$$

---

* Presented to the American Mathematical Society October 27, 1900. Received for publication February 23, 1901.

† *On surfaces enveloped by spheres belonging to a linear spherical complex*, vol. 1 (1900), p. 371.

(3)* $$r' = \frac{r_0\kappa_0(x^2 + y^2 + z^2 - r_0^2)}{x^2 + y^2 + z^2 - r_0^2\kappa_0^2}.$$

The spheres S in question are subjected to a single condition. The elimination of x, y, z from (2) and (3) gives

$$x'^2 + y'^2 + z'^2 - r_0^2 - r'^2 - 2r_0r'\kappa = 0,$$

if

$$\kappa = \frac{1 + \kappa_0^2}{2\kappa_0}.$$

The ∞^3 spheres S, therefore, belong to a linear spherical complex A whose fundamental sphere is σ and whose "constant" equals κ.

Special features of Γ are the following:—Points on σ correspond to themselves. Points on the sphere $\sigma_0(0, 0, 0, -r_0\kappa_0)$ give the ∞^2 planes of A, and these all touch the sphere $\sigma_0'(0, 0, 0, -r_0\kappa)$. The sphere σ_0 belongs to A. *Any* point of R_3 at infinity and *not* on the imaginary circle C corresponds to σ_0, while a point P of C gives the ∞^1 spheres of radius $-r_0\kappa_0$ whose centers lie upon the line joining P with the origin. Therefore the *fundamental configurations* of Γ are the points of C in R_3 and the sphere σ_0 of A.

A property of the correspondence of prime importance is established by the equation

(4)
$$(x_1' - x_2')^2 + (y_1' - y_2')^2 + (z_1' - z_2')^2 - (r_1' - r_2')^2 = \frac{r_0^2(1 - \kappa_0^2)[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2]}{(x_1^2 + y_1^2 + z_1^2 - r_0^2\kappa_0^2)(x_2^2 + y_2^2 + z_2^2 - r_0^2\kappa_0^2)},$$

in which $S_1(x_1', y_1', z_1', r_1')$ and $S_2(x_2', y_2', z_2', r_2')$ are taken as corresponding to $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, respectively. This gives

THEOREM 1. *The ∞^1 points P of a minimum line intersecting σ in P_0 become, under Γ, ∞^1 spheres of A touching one another at P_0.*

By considering the minimum cone with vertex P, we easily get a second result:

A sphere S of a linear spherical complex A is touched by ∞^2 spheres of A along its circle of intersection with the fundamental sphere.†

This circle we call the *trajectory circle* of S.

Let E be any plane passing through P, and let the two minimum lines drawn in E through P intersect σ in p_1, p_2. Then p_1, p_2 are vertices of a skew quadrilateral $p_1p_2P_1P_2$ of minimum lines lying in S, while p_1p_2, P_1P_2 are polar lines with respect to S, since the planes $p_1P_1p_2$, $p_1P_2p_2$ touch S at P_1 and P_2 respectively. Denoting these planes by E_1, E_2, we have the result:

*Cf. Transactions, p. 383. For $\kappa_0 = 0$, the correspondence is the usual inversion in σ. Evidently the case $\kappa_0^2 = 1$ is of no interest.

†LIE, Mathematische Annalen, vol. 5, p. 207.

THEOREM 2. *A surface element* (P, E) *corresponds under* Γ *to two elements* (P_1, E_1), (P_2, E_2) *of the sphere* S. *The planes* E, E_1, E_2 *intersect in the radical plane of* S *and* σ, *while the circle drawn through* P, P_1, P_2 *is orthogonal to* E, S, σ.

For example, if P is at infinity in E and not on the imaginary circle, then (P_1, E_1) and (P_2, E_2) are on σ_0 while E_1, E_2, E are parallel.

If, now, P moves along the minimum line Pp_1, then P_1 and P_2 are displaced on p_1P_1 and p_1P_2 respectively, this being true for any plane E through Pp_1. Therefore,

A minimum line regarded as defined by its ∞^2 *surface elements becomes under* Γ *two minimum lines whose plane belongs to* A, *the three lines intersecting on* σ.

The result just stated is of fundamental importance, for we may at once deduce the following properties:

A sphere s becomes under Γ two spheres s_1 and s_2, and s_1, s_2, s intersect on σ. If s degenerates into a plane, then s_1 and s_2 touch the sphere σ_0.

The spheres determined by $P_1p_1p_2P$ and $P_2p_1p_2P$ are invariant. Hence the planes tangent to either along its intersection with σ must be invariant, i. e., must touch the sphere σ_0. The conclusion is:

THEOREM 3. *The spheres of a linear spherical complex* A_0 *with fundamental sphere* σ *and constant* κ_0 *are invariant under* Γ. *The* ∞^2 *spheres of* A_0 *tangent to a given sphere* s *also touch each of two spheres* s_1 *and* s_2, *into which* s *goes over by* Γ.

From the properties enunciated in this last theorem, I have called the transformation of space which Γ establishes *inversion in the complex* A_0,* and have symbolized it by (A_0).

Theorem 2 may be given another interesting form. For if we denote the two spherical images of the plane E on the sphere S by p_0 and p_0', i. e., the line joining p_0 and p_0' is that diameter of S which is perpendicular to E, then the spherical inversion in the center P which leaves S invariant transforms p_0 and p_0' into P_1 and P_2 respectively. Thus we have the result:

By the spherical inversion with center P *which transforms* S *into itself, the spherical image on* S *of any surface element* (P, E) *inverts into the point* P_1 *of the corresponding surface element* (P_1, E_1).

Next consider all surface elements with common P whose planes intersect in a line L. The special representation of these elements on S gives the points of the great circle C whose axis is parallel to L. Then the transformed elements lie along the small circle c on S into which C inverts in the center P. The circle c intersects the trajectory circle of S orthogonally, since its plane passes through the pole of the radical plane of P and S with respect to S.

*Cf. Transactions, vol. 1, p. 376.

The results established in this section hold with certain obvious exceptions for any number of dimensions, although the reasoning here given has been restricted to the case of ordinary space.

§ 2. *The corresponding contact transformation.*

We proceed now to find the equations of (A_0) as a contact transformation.

For convenience, we write in (1) x_1, x_2, x_3 for x, y, z, and similarly for X, Y, Z. As coördinates of the surface element (P, E) we may take

$$x_1, x_2, x_3; \quad p_1, p_2, p_3, \quad \text{or} \quad (x, p)$$

if $p_1 : p_2 : p_3$ determine the direction of the normal to E. Let (P, E) become (X, P) by (A_0); then we have to determine X and P from the equations †

$$\Phi = 0, \quad \rho p_i - \frac{\partial \Phi}{\partial x_i} = 0, \quad \rho P_i + \frac{\partial \Phi}{\partial X_i} = 0 \qquad (i = 1, 2, 3), \tag{5}$$

after elimination of ρ.

The result is

$$\left.\begin{aligned} X_i &= x_i - \frac{(x_i s - p_i \delta)}{s(u + \delta^2) - 2\delta l}(u - r_0^2), \\ P_i &= \frac{p_i(u - \delta^2) - 2x_i(l - \delta s)}{r_0^2 - \delta^2} \qquad (i = 1, 2, 3), \end{aligned}\right\} \tag{6}*$$

where

$$u = \sum_i x_i^2, \quad s^2 = \sum_i p_i^2, \quad l = \sum_i p_i x_i, \quad \delta = -r_0 k_0.$$

To these may be added, if $S^2 = \sum_i P_i^2$,

$$S = \frac{s(u + \delta^2) - 2\delta l}{r_0^2 - \delta^2}. \tag{7}$$

The equation $s(u + \delta^2) - 2\delta l = 0$ defines the congruence of spheres of radius δ and with centers on the minimum cone whose vertex is the center of σ.

The element (P, E) is invariant when, and only when, P is on $\sigma\,(u = r_0^2)$ and E is tangent to $\sigma_0 (l - \delta s = 0)$. This result shows that the spheres of the complex A_0 are unchanged by (A_0) as already stated.

The following special results are to be noted, and may be readily verified.

An element (P, E) such that P is on σ and E belongs to A transforms into ∞^1 (P', E')'s, where E' is a minimum plane, and P' *any* point on the minimum line in E' through P.

An element (P, E) belonging to σ_0 gives the ∞^1 elements of E at infinity.

---

* Cf. LIE-ENGEL, *Transformationsgruppen*, vol. II, p. 151.

† Transactions, loc. cit., p. 385. As remarked above, $\delta = 0$ gives inversion in σ.

An element (P, E) of the imaginary circle at infinity goes over into the ∞^2 elements of the sphere into which P inverts under (A_0) (§ 1).

§ 3. *Duality of the geometry of reciprocal radii and geometry within a linear spherical complex.*

The group of the geometry within a linear spherical complex A is that subgroup G_{10} of the general G_{15} of all sphere-sphere contact transformations under which A is invariant. Any transformation of G_{10} is compounded of inversions in complexes in involution with A. The contact transformation (A_0) which carries A over into the complex of all points transforms G_{10} into the group of the geometry of reciprocal radii, every transformation of which is compounded of spherical inversions. Thus within the general higher spherical geometry, *the two geometries mentioned appear equally justified.* This result is fundamental for what follows. Let us now assign to the infinite region the characteristics demanded by the geometry of reciprocal radii. Then the points on the imaginary circle are to be excluded and the infinite region becomes the "point at infinity." The fundamental configurations of the correspondence Γ (§ 1) disappear, while for (A_0) the $\infty^2(P, E)$'s whose points are on σ and whose planes touch σ_0' constitute the fundamental configuration. These go over into the ∞^3 minimum elements (finite).

We may now mention a number of the more interesting correspondences in the two geometries; in each case the first statement refers to the geometry of reciprocal radii.

Surface f, point locus.	Surface F, enveloped by ∞^2 spheres of A.
Curve of intersection c of f and σ.	Double curve c of F on σ.
Curve of intersection c' of f and σ_0.	Developable of bitangent planes to F, enveloping σ_0'.
Principal spheres of f.	Principal spheres of F.
Minimum developable intersecting σ in c_0.	Curve c_0.
Singular line of curvature on f, viz., the curve of contact of f with the circumscribing minimum developable f'.	Line of curvature on F lying on σ, along which the principal spheres have four-point contact; this line of curvature and the double curve c above mentioned constitute the entire intersection of F and σ.*

* This singular line of curvature on any surface f is determined by differentiation and elimination. These operations suffice also for the determination of the line of curvature in question on F.

Confocal surfaces, i. e., surfaces inscribed in a minimum developable f'.	Surfaces enveloped by ∞^2 spheres of A and inscribed in a developable, whose planes belong to A, along a line of curvature on σ.
Lines of curvature on f.	Lines of curvature on F.
Minimum lines on f.	Minimum lines on F.

The congruence of ∞^2 circles orthogonal at once to f and the fundamental sphere σ intersect F also orthogonally, such that each circle cuts F in the two points of contact of an enveloping sphere (§ 3). * The ∞^1 circles of this congruence intersecting f along any curve trace out upon F the corresponding curve.

It will be remarked that the developables whose planes belong to A, i. e., touch the sphere σ_0', play an analogous rôle in the geometry here discussed to that of minimum developables in the geometry of reciprocal radii. The first mentioned surfaces are remarkable for the fact that their lines of curvature (non-linear) lie upon spheres concentric with the fundamental sphere.

Continuing the discussion of the duality, we have as corresponding:

Curve C, point locus.	Surface F enveloped by ∞^1 spheres of A, i. e., an annular surface.
Point of intersection of C and σ.	Nodes on F.
Points of intersection of C and σ_0.	Planes touching F along circles.
Minimum developable passed through C and intersecting σ in c.	Line of curvature c on F, along which the principal spheres have four-point contact, the complete insection of F and σ.
Minimum curve, cuspidal edge of minimum developable intersecting σ in c.	Curve c.

The locus of centers (x', y', z') of the spheres of A enveloping F is found from f by the *point* transformation defined by equations (2), § 1.*

The conformal property of transformations of the group of the geometry of reciprocal radii is expressed in the geometry of surface elements as follows:— (P, E) and (P, E') go over into (P', E_1) and (P', E_1') such that the angle of EE' equals that of E_1E_1'. For definiteness orientation must be assigned to the planes in question. Now by inversion in the complex (A_0) which transforms P into S, (P, E) and (P, E') go over into the elements of S, $(P_1 E_1)$ and (P_2, E_2), and P_1 and P_2 lie on a circle of S orthogonal to the trajectory circle

* For a discussion of the locus of centers, see Transactions, vol. 1, p. 383, seq.

(§ 1), while the minimum elements of the pencil (P, E), (P, E'), invert into the elements common to these two circles. We have then from the results of the last part of § 1 the

THEOREM. *If the linear spherical complex A invert by (A_0) into the complex of all points so that (P, E), (P, E') become (P_1, E_1), (P_1', E_1') on the sphere S, and if the circle on S through P_1 and P_1' orthogonal to the trajectory circle intersect the latter in P_0 and P_0', then the logarithm of the cross ratio $(P_1P_0, P_1'P_0')$ divided by $2\surd-1$ equals one-half the angle of the planes E and E', and this cross ratio is unaltered by every transformation of G_{10}.*

Let, now, a curve C invert in (A_0) into the annular surface F enveloped by ∞^1 spheres of A. Any curvature strip* along C becomes under (A_0) a curvature strip on F, say along the curve c on F. By the theorem of JOACHIMSTHAL, any strip along C intersecting that mentioned under a constant angle is also a curvature strip, therefore inverts under (A_0) into a curvature strip on F, and in this way *all* non-circular lines of curvatures c on F are found. Hence the result:

Theorem of Joachimsthal. Two curvature strips along a curve intersect under a constant angle.

The non-circular lines of curvature of an annular surface whose spheres belong to a linear spherical complex determine projective ranges on the circles of curvature.†

In the same way we get as dual theorems:

Theorem of Dupin. Any system of ∞^3 surfaces mutually orthogonal intersect along lines of curvature.

If in any system of ∞^3 surfaces enveloped by the spheres of a linear spherical complex A the six points of contact of every sphere of the complex are the intersections of three mutually orthogonal circles on that sphere whose common orthogonal circle is the trajectory circle, then the annular surface enveloped by ∞^1 spheres of A and tangent to a surface of the system along a line of curvature will touch also a second surface of the system along a line of curvature.

* i. e., a system of ∞^1 elements belonging to C whose normals form a developable surface.

† This theorem holds for all annular surfaces. But when the enveloping spheres belong to a linear spherical complex the non-circular lines of curvature are determined by *quadratures*, for a *particular* solution of the corresponding RICCATI equation is known, viz., the line of curvature c above mentioned.

Other theorems of interest might be added to those already given, but enough has been given to justify the study of the duality.

It is evident that the general surface within a linear spherical complex will intersect the plane at infinity only along the imaginary circle. Furthermore, such a surface will contain no *finite* minimum element, whence the result:

A general algebraic surface in the geometry within a general linear spherical complex is of equal order and class, cutting the plane at infinity only in the imaginary circle, and the enveloping minimum developable is singular.

§ 4. *The group of inversions in all linear spherical complexes with common fundamental sphere.*

If (A) and (A') are inversions in complexes A and A' having a common fundamental sphere σ, then *

$$(AA') \equiv (\Omega A''),$$

where (Ω) denotes change in orientation and A'' also has σ for fundamental sphere. If $\kappa, \kappa', \kappa''$ are the constants of A, A', A'' respectively, then

$$\kappa'' = \pm \frac{1 - \kappa\kappa'}{\kappa - \kappa'}. \tag{8}$$

Now (Ω) results from (6) and (7) when $\delta = \infty$, for then $(x; p)$ becomes $(x; p)$ *but* s *changes sign.* Hence (Ω) plays here the rôle of the identical transformation, and we have

THEOREM 4. *The equations* (6) *define a group* G_1 *of contact transformations in the parameter* $\kappa_0 (\delta = -r_0\kappa_0)$. *The composition law is given by* (8).

We find the infinitesimal transformation $U(f)$ of G_1 in the usual way by setting $\delta = \infty + \delta t$, and obtain

$$\frac{\delta x_i}{\delta t} = \frac{p_i}{s}(u - r_0^2), \qquad \frac{\delta p_i}{\delta t} = -2x_i s \qquad (i = 1, 2, 3);$$

therefore G_1 is generated by

$$U(f) \equiv \frac{u - r_0^2}{s}\left(p_1\frac{\partial f}{\partial x_1} + p_2\frac{\partial f}{\partial x_2} + p_3\frac{\partial f}{\partial x_3}\right) - 2s\left(x_1\frac{\partial f}{\partial p_1} + x_2\frac{\partial f}{\partial p_2} + x_3\frac{\partial f}{\partial p_3}\right). \tag{9}$$

We determine now the *invariants* of G_1. The equation $U(f) = 0$ has the six solutions

* This follows from Theorem 4, p. 378, Transactions, vol. 1 (1900).

$$(10)\qquad \begin{cases} a_3 \equiv x_1p_2 - x_2p_1, \quad a_1 \equiv x_2p_3 - x_3p_2, \quad a_2 \equiv x_3p_1 - x_1p_3, \\ \beta_i \equiv x_i l - \tfrac{1}{2}\, p_i(u + r_0^2) \qquad (i = 1, 2, 3), \end{cases}$$

connected by the identical relation

$$(11)\qquad a_1\beta_1 + a_2\beta_2 + a_3\beta_3 \equiv 0,$$

but otherwise independent. Hence

THEOREM 5. *Every function invariant under G_1 has the form*

$$F(a_1, a_2, a_3, \beta_1, \beta_2, \beta_3),$$

where F is homogeneous of dimension zero.

The a, β have an interesting geometrical significance. For, if L is the line of intersection of the plane E of (P, E) $[(x; p)]$ with the radical plane of P and the fundamental sphere σ, then the equations of L in current coördinates X_1, X_2, X_3 are

$$a_3X_1 = a_1X_3 + \beta_2, \quad a_3X_2 = a_2X_3 + \beta_1,$$

i. e., the a, β are the *line coördinates* of L. The plane through P perpendicular to L is

$$a_1X_1 + a_2X_2 + a_3X_3 = 0;$$

and if C be its intersection with L, then the coördinates of C are

$$\frac{a_2\beta_3 - a_3\beta_2}{a_1^2 + a_2^2 + a_3^2}, \quad \text{etc.},$$

and

$$CP^2 = \frac{\sum_i \beta_i^2}{\sum_i a_i^2} - r_0^2.$$

These results give

THEOREM 6. *By the transformations of G_1 upon any surface element (P, E) the point P is made to describe the circle orthogonal at once to E and the fundamental sphere, while E turns round the axis of this circle.*

For the coördinates of the axis, the center, and the radius of the circle in question are invariants.

The most general *system* of ∞^4 elements invariant under G_1 will be defined by an equation of the form

$$(12)\qquad F(a_1, a_2, a_3, \beta_1, \beta_2, \beta_3) = 0,$$

F being homogeneous.

This equation, however, also defines a line complex, the locus of ∞^3 axes $(a; \beta)$ of the circles of Theorem 5. We are therefore led to a solution of the problem:—*To determine all partial differential equations of the first order*

the integral surfaces of which cut orthogonally a given system of ∞^3 *circles orthogonal to a given sphere* σ.

For the given system of circles is uniquely defined by the line complex formed of their axes, (a, β) i. e., by an equation (11). If then we replace the a and β from equations (10) the problem is solved. *

The "characteristic function" of the infinitesimal homogeneous contact transformation $U(f)$ is

$$H \equiv s(u - r_0^2),$$

since †

$$U(f) \equiv \sum_{i=1}^{3} \left(\frac{\partial H}{\partial p_i} \frac{\partial f}{\partial x_i} - \frac{\partial H}{\partial x_i} \frac{\partial f}{\partial p_i} \right).$$

From theorem 5 we see that the ∞^1 surfaces resulting from transformation of an *arbitrary* surface by the group G_1 belong, by the theorem of RIBAUCOUR, to an orthogonal system. Furthermore, under $U(f)$ the lines of curvature go over into lines of curvature. Finally, SOPHUS LIE has found that the most general infinitesimal contact transformation possessing these properties ‡ has for characteristic function

$$H \equiv s(a_0 u + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4),$$

into which $U(f)$ transforms by a translation.

§ 5. *Generalization to space of n dimensions; in particular, to* $n = 2$.

The preceding discussion may readily be generalized to n dimensions. Theorems 1–5 hold, with trivial changes, in general.

Confining ourselves for brevity to the discussion of the problem mentioned in § 4 for the case $n = 2$, we may state the result:

The most general differential equation of the first order whose integral curves cut orthogonally a given system of ∞^1 circles orthogonal to a given circle σ is determined thus:—

The center of σ being the origin and its radius r_0, let the given system of circles be defined by the locus of centers

$$F(a, \beta, \gamma) = 0,$$

F being homogeneous in the homogeneous rectangular coördinates a, β, γ. Then setting

---

* Writing, of course,

$$p = \frac{\partial x_3}{\partial x_1} = -\frac{p_1}{p_3}, \quad q = \frac{\partial x_3}{\partial x_2} = -\frac{p_2}{p_3}.$$

The integration problem of such a differential equation is much simplified by observing that its integral surfaces admit of $U(f)$.

† Cf. LIE-ENGEL, *Transformationsgruppen*, vol. 2, p. 263.

‡ Vide Leipziger Berichte, 1889, p. 153. Cf. also DARBOUX, *Leçons sur les systèmes orthogonaux*, vol. 1, pp. 56, 72, in which (A_0) is called the *transformation de Ribaucour*.

$$\gamma = x_1 p_2 - x_2 p_1, \quad a = \tfrac{1}{2} p_2 (u + r_0^2) - x_2 l, \quad \beta = x_1 l - \tfrac{1}{2} p_1 (u + r_0^2),$$

$$y' = \frac{dx_2}{dx_1} = -\frac{p_1}{p_2},$$

where

$$u = x_1^2 + x_2^2, \quad l = p_1 x_1 + p_2 x_2,$$

we obtain the required equation,

$$F\left(\tfrac{1}{2}(x^2 + y^2 + r_0^2) + y(xy' - y),\ \frac{y'}{2}(x^2 + y^2 + r_0^2) - x(xy' - y),\ x + yy'\right) = 0,$$

wherein x and y have been written for x_1 and x_2, and F is an *arbitrary homogeneous function.*

The characteristic function for the infinitesimal contact transformation $U(f)$ now takes the form*

$$W \equiv -(x^2 + y^2 - r_0^2)\sqrt{1 + y'^2}.$$

§ 6. *Some special cases of the preceding results.*

Inversion in a general linear spherical complex is transformed by a spherical inversion whose center lies on the fundamental sphere into inversion in a plane spherical complex, i. e., into the transformation of LAGUERRE.† The fundamental sphere σ is replaced by a fundamental plane π. If π be taken for $z = 0$, the point-sphere correspondence Γ of § 1 becomes

$$(13) \qquad \Phi' \equiv (1 - \kappa_0^2)\{(X - x)^2 + (Y - y)^2 + (Z - z)^2\} + 4zZ = 0,$$

and is *the most general involutory point-sphere correspondence such that a point P and its associated sphere S have a common radical plane.*

The ∞^3 spheres S belong to a plane complex E with fundamental plane π and constant $\kappa = 1 + \kappa_0^2/2\kappa_0$, as before.

Characteristic properties of the transformation are (1), planes transform into planes, the plane at infinity being invariant, and (2), the distance of two points P and P' equals the length of a common tangent of the corresponding spheres.

Theorems 1, 2, 3 hold here.

The equations of the contact transformation, i. e., inversion in the complex E_0 with fundamental plane $z = 0$ and constant κ_0 are readily found as before (§ 2). As *fundamental configurations* of (E_0) appear (1), the ∞^2 surface elements whose points lie in π and whose planes make with π the angle $\cos^{-1}\kappa$, and (2), the ∞^2 elements of the imaginary circle at infinity. An element of (1) transforms into the ∞^1 minimum elements along a minimum line, while to an

* Cf. SOPHUS LIE, *Geometrie der Berührungstransformationen*, vol. 1, p. 150.

† SMITH, Annals of Mathematics, July, 1900.

element of (2) correspond the ∞^2 elements of a plane belonging to E. Especially noteworthy is the fact that two elements (P_1, E), (P_2, E) having the same plane E go over into elements of the same sort, (P_1', E'), (P_2', E') such that $P_1P_2 = P_1'P_2'$. This property is expressed thus:

Tangential distance is invariant under inversion in a plane spherical complex.

By a spherical inversion elements having a common point (P, E_1), (P, E_2) transform into elements of like nature (P', E_1'), (P', E_2'), and such that the angle of E_1 and E_2 equals that between E_1' and E_2'. I have elsewhere bespoken for the transformation of LAGUERRE an attention it has not as yet received,* and the analogy just remarked would seem to warrant this.

Within the general G_{15} (§ 3) of all sphere-sphere contact transformations in space is contained as subgroup a G_{10}' of all such transformations as are compounded of inversions in plane spherical complexes. This G_{10}' is the group of the geometry of reciprocal directions. The contact transformation (E_0') inverts the group G_6 of all transformations compounded of plane reflections into that subgroup G_6' of G_{10}' under which the plane spherical complex E remains invariant. In this way we establish a *duality of the Euclidean geometry* (excluding similitude) *to spherical geometry within a plane spherical complex.*

It is not proposed to develop this question. There is of course much in § 3 which is valid here, but we have in addition the very interesting correspondence of distance to tangential distance.

The group of ∞^1 contact transformations made up of inversions in all plane complexes having the same fundamental plane $x_3 = 0$ is generated by the infinitesimal contact transformation $U(f)$ whose characteristic function is

$$H \equiv 2x_3 s,$$

i. e.,

$$U(f) = \frac{2x_3}{s}\left(p_1\frac{\partial f}{\partial x_1} + p_2\frac{\partial f}{\partial x_2} + p_3\frac{\partial f}{\partial x_3}\right) - 2s\frac{\partial f}{\partial p_3}.$$

The general invariant function has the form

$$F(p_1, p_2, p_1x_1 + p_2x_2 + p_3x_3, x_1p_2 - x_2p_1, x_3s),$$

F being homogeneous of dimension zero, while every invariant system of ∞^4 elements is defined by setting any homogeneous function of these arguments equal to zero.

Under the present G_1 carried out on the surface element (P, E) the point P describes a circle c whose center lies in the fundamental plane while E turns around the axis l of this circle.

The center of c is

* In the Annals of Mathematics, July, 1900.

$$\left(\frac{p_1 l + p_2(x_1 p_2 - x_2 p_1)}{p_1^2 + p_2^2}, \quad \frac{p_2 l - p_1(x_1 p_2 + x_2 p_1)}{p_1^2 + p_2^2}\right),$$

and its radius equals $x_3 s\, p_1 p_2 (p_1^2 + p_2^2)^{-\frac{1}{2}}$. The axis l intercepts on OX and OY the distances

$$\frac{l}{p_1}, \quad \frac{l}{p_2}.$$

Finally, the preceding discussion put us in a position to solve the problem:

To determine all partial differential equations of the first order whose integral surfaces intersect orthogonally a given system of ∞^3 circles whose centers lie in a given plane and whose planes are perpendicular to that plane.

If now the fundamental plane of Γ is the plane at infinity, the correspondence becomes

$$(X - x)^2 + (Y - y)^2 + (Z - z)^2 = a^2.$$

The corresponding contact transformation is a dilatation of a point P into the sphere S of constant radius a and center at P, while the problem mentioned above becomes this: *To determine all partial differential equations of the first order whose integral surfaces have for normals the lines of a given line complex.**

§ 7. *Geometrical definition of the correspondence* Γ.

The subject under discussion appears in a new light from a definition of the fundamental correspondence (§ 1) now to be established.

Let P be any point in space R_n, and O a point in R_{n+1} external to R_n. Then by a spherical inversion in the center O the minimum cone V of R_{n+1} whose vertex is P transforms into a minimum cone V' of vertex P'. The cone V' intersects R_n in a sphere S, while the sphere S_0 into which R_n inverts cuts R_n in a sphere σ. Also V and S intersect on σ. Reciprocally, the minimum cone in R_{n+1} with vertex any point P'' on S, passes through P', and hence inverts into a minimum cone containing P. We see, therefore, that the correspondence of P and S possesses the characteristics of Γ. This gives

THEOREM 7. *The point-sphere correspondence Γ in R_n may be thus defined: A spherical inversion in a center in R_{n+1} external to R_n inverts the minimum cone of R_{n+1} whose vertex is a point P of R_n into a minimum cone whose intersection with R_n is the corresponding sphere S.*

To proceed to the other properties of Γ, we choose the sphere as element of R_n. The ∞^{n+1} points P' of R_{n+1} project by *minimal projection* into the ∞^{n+1} spheres S of R_n. In fact, the minimum cone in R_{n+1} with vertex P' intersects R_n in the sphere S. We shall lose nothing in generality by confining the discussion to $n = 2$; then R_2 becomes a plane π; and the ∞^3 points

*Cf. SOPHUS LIE, Mathematische Annalen, vol. 5 (1872), p. 201.

P' of space project into the ∞^3 circles C of π, such that the center of C is the orthogonal projection of P' on π while the radius of C equals the ordinate of P' relative to C multiplied by $\sqrt{-1}$.*

The ∞^2 points of any surface F in R_3 project into the ∞^2 spheres of a circular complex whose curve of singularities is the intersection with π of the minimum developable circumscribed to F. In particular, let F be a sphere S_0. If then S_0 cuts π in the circle σ of radius r_0 and center O, a simple calculation shows that any point on S_0 projects into a circle in π of center c' and radius r' such that

$$\overline{cc'}^2 - r_0^2 - r'^2 - 2r_0r'\kappa = 0,$$

where

$$R_0^2 = r_0^2(1 - \kappa^2),$$

R_0 being the radius of S_0. We see then

THEOREM 8. *The points of a sphere S_0 in R_{n+1} project by minimal-projection on R_n into the spheres of a linear spherical complex whose fundamental sphere is the intersection of S_0 and R_n.*

Spherical inversion in the sphere S_0 in R_3 goes over into inversion in the corresponding circular complex in π. Hence the conformal group of point transformations in R_3 transforms by minimal projection into the group of all circular contact transformations in π. A plane reflection in R_3 becomes inversion in a line complex, which, if the plane of the reflexion is parallel to π, reduces to a dilatation. The theorems proved in §§ 2, 3 in the paper already quoted † are readily verified by the known properties of inversions in space.

If now we consider either of the two commutative spherical inversions (I_0) in space by which π transforms into the sphere S_0, then to (I_0) corresponds in π by the minimal-projection, the inversion in the circular complex A_0 to which Γ gives rise as a contact transformation. The main results of the previous sections are thus easily established.

Let C be any curve on S_0, and denote the minimum developable which is the envelope of the minimum cone V whose vertex describes C by F'. Then F' intersects π in a curve c which is enveloped by ∞^1 circles in the circular complex A. The locus of centers of the circles enveloping c is the orthogonal projection of C on π, and c is derived by inversion in a circular complex A_0 from a primitive curve c', regarded as point locus, which is the stereographic projection of C on π. In particular, if C is a general sphero-quartic, the minimum developable in question envelops a system of confocal cyclides, and is of order 16 passing 8 times through the imaginary

* Cf. for further details, LIE, *Geometrie der Berührungstransformationen*, vol. 1, p. 437, or KLEIN, *Höhere Geometrie*, p. 472.

† SMITH, Transactions, vol. 1 (1900).

circle at infinity and containing five double sphero-quartics of which C is one. The curve c in π is therefore the curve of singularities of the general quadratic circular complex, and the properties enumerated in the paper already quoted * follow at once.

We may now state

THEOREM 9. *Given an M_r on a sphere S_0 in R_{n+1} and let the manifoldness which is the envelope of a minimum cone whose vertex describes M_r intersect R_n in M'_{n-1}. Then M'_{n-1} is enveloped by ∞^r spheres of a linear spherical complex in R_n, this complex being the locus of the spheres into which the points of S_0 pass over by minimal projection. The locus of centers of the spheres enveloping M'_{n-1} is the orthogonal projection of M_r on R_n, and M'_{n-1} may be derived by inversion in a linear spherical complex from a primitive m_r regarded as point locus, viz. the stereographic projection of M_r on R_n.*

SHEFFIELD SCIENTIFIC SCHOOL, NEW HAVEN, CONN., *February*, 1901.

*SMITH, Transactions, vol. 1, p. 390.

A NEW DETERMINATION OF THE PRIMITIVE CONTINUOUS GROUPS IN TWO VARIABLES*

BY

H. F. BLICHFELDT

The primitive continuous groups of point-transformations in two variables can, by a proper choice of the variables, be transformed into *projective groups* of the plane, a result LIE obtains after determining the canonical forms of the primitive groups.† This fact can, however, be established from the general properties of such groups, and its use leads to a new determination of these primitive groups, to show which is the object of this paper.

A primitive group will be defined as a group which does not leave invariant a differential equation of the first order.‡ Such a group is at least three-parametric, as a two-parametric group possesses a differential invariant of the first order, J say, and therefore an invariant differential equation of the first order, $f(J) = \text{constant}$.§

§ 1.

It will first be necessary to show that *any group in two variables and of more than two parameters leaves invariant at least one differential equation, which is integral and algebraic in the derivatives* dy/dx, d^2y/dx^2, *etc., involved.*

Let $Xf \equiv \xi\, \partial f/\partial x + \eta\, \partial f/\partial y$ indicate any one of the infinitesimal transformations of the group. If we "extend" these n times,‖ and write $y_1, y_2, \cdots, y_n$ for $dy/dx, d^2y/dx^2, \cdots, d^ny/dx^n$, and $p, q, p_1, p_2, \cdots, p_n$ for $\partial f/\partial x, \partial f/\partial y, \partial f/\partial y_1, \partial f/\partial y_2, \cdots, \partial f/\partial y_n$, respectively, we easily find these extended transformations to be of the form

$$\xi p + \eta q + [\eta_x + (\eta_y - \xi_x)y_1 - \xi_y y_1^2]p_1 + (P_2 y_2 + Q_2)p_2 + \cdots + (P_n y_n + Q_n)p_n,$$

* Presented to the Society April 27, 1901. Received for publication March 3, 1901.

† S. LIE: *Vorlesungen über continuierliche Gruppen*, herausgegeben von Dr. SCHEFFERS, p. 359. This book will hereafter be designated *Continuierliche Gruppen*.

‡ *Invariant differential equation* defined in *Continuierliche Gruppen*, pp. 213–214, and in *Differentialgleichungen mit bekannten infinitesimalen Transformationen*, by S. LIE, p. 277.

§ *Continuierliche Gruppen*, p. 228.

‖ *Continuierliche Gruppen*,, pp. 213–214.

where

$$P_m = -(m+1)\xi_y y_1 - m\xi_x + \eta_y,$$

$$Q_2 = -\xi_{yy}y_1^3 + (-2\xi_{xy} + \eta_{yy})y_1^2 + (-\xi_{xx} + 2\eta_{xy})y_1 + \eta_{xx},$$

$$Q_3 = -3\xi_y y_2^2 + y_2 R_3(x, y, y_1) + S_3(x, y, y_1),$$

$$Q_m = -\tfrac{1}{2}m(m-1)\,\xi_y y_{m-1} + y_{m-1}R_m(x, y, y_1) + S_m(x, y, y_1, \cdots, y_{m-2}) \quad (m>3),$$

the quantities R and S being integral algebraic functions of $y_1, y_2, \cdots, y_{m-2}$.

We shall indicate the n times extended infinitesimal transformation Xf by

$$X^n f \equiv \xi p + \eta q + \sum_{i=1}^{n} \eta^{(i)} p_i.$$

LIE has proved that an m-parametric group possesses one, and only one, differential invariant, J, of order less than m.*

First, let J be of order $m-1$. Then must the m equations

$$X^{m-2}f \equiv \xi \frac{\partial f}{\partial x} + \eta \frac{\partial f}{\partial y} + \sum_{i=1}^{m-2} \eta^{(i)} \frac{\partial f}{\partial y_i} = 0 \tag{1}$$

be independent of each other, $\xi\,\partial f/\partial x + \eta\,\partial f/\partial y$ being a type of the m independent infinitesimal transformations of the group. For, were they not independent, we should have a complete system of fewer than m equations in the m variables $x, y, y_1, \cdots, y_{m-2}$, which would have as solution a differential invariant of degree $m-2$ at most, contrary to hypothesis.

The determinant D of the m^2 coefficients ξ, η, $\eta^{(i)}$ of the equations (1) will then not vanish identically. If it contains any one of the variables $y_1, y_2, \cdots, y_{m-2}$, then will $D=0$ be an invariant differential equation,† and by examining the form of the quantities $\eta^{(i)}$ given above, we find that D must be an integral algebraic function of $y_1, y_2, \cdots, y_{m-2}$.

The determinant D may be a function of x, y only, say $\phi(x, y)$, which is easily seen to be the determinant of the m^2 quantities $\xi, \eta, \eta_x, \eta_{xx}, \cdots, \partial^{m-2}\eta/\partial x^{m-2}$.‡ Now, by choosing y for the independent variable instead of x, the transformed determinant D of (1) will be seen to contain $\pm\phi\,(dx/dy)^{(m+1)(m-2)/2}$ as the highest power of dx/dy. In this case then, *if* $m>2$, $D=0$ will be an invariant differential equation, integral and algebraic in the derivatives involved.

---

* *Continuierliche Gruppen*, page 230. In the text J is given by J_{m-1}.

†S. LIE, *Theorie der Transformationsgruppen*, bearbeitet unter Mitwirkung von Dr. F. ENGEL; I. Abschnitt, Kapitel 14.

‡ These quantities are the terms of $X^{m-2}f$ free from the derivatives $y_1, y^2, \cdots, y_{m-2}$.

Next, let J be of less than the $(m-1)$-th order. Every differential equation $J=$ constant is invariant, and therefore the equation $dJ\,dx=0$ is invariant and is of the nth order, say, where $n \leqq m-1$.

The various equations:

$$(2) \qquad X^n\left(\frac{dJ}{dx}\right) \equiv \xi\frac{\partial}{\partial x}\left(\frac{dJ}{dx}\right)+\eta\frac{\partial}{\partial y}\left(\frac{dJ}{dx}\right)+\sum_{i=1}^{n}\eta^{(i)}\frac{\partial}{\partial y_i}\left(\frac{dJ}{dx}\right)=0$$

cannot be identities, since dJ/dx is not a differential invariant, there being only one such of order less than m, which is here J. Since this is of the $(n-1)$-th order, only n of the equations

$$\xi\frac{\partial f}{\partial x}+\eta\frac{\partial f}{\partial y}+\sum_{i=1}^{n-1}\eta^{(i)}\frac{\partial f}{\partial y_i}=0$$

can be independent, and it is therefore possible to eliminate the quantities

$$\frac{\partial}{\partial x}\left(\frac{dJ}{dx}\right),\quad \frac{\partial}{\partial y}\left(\frac{dJ}{dy}\right),\quad \frac{\partial}{\partial y_i}\left(\frac{dJ}{dx}\right) \qquad (i=1,2,\cdots,n-1),$$

in various ways from (2) and thus to obtain equations of the form

$$M\frac{\partial}{\partial y_n}\left(\frac{dJ}{dx}\right)=0,$$

where the quantities M are determinants of the coefficients $\xi, \eta, \eta', \cdots, \eta^{(n)}$ and are not all zero identically, but must evidently all vanish on account of $dJ\,dx=0$. Since the determinants M are integral and algebraic in $y_1, y_2, \cdots, y_n$, this must be the case with $dJ/dx=0$, after a possible reduction.

In all cases then, we have an invariant differential equation, integral and algebraic in the derivatives involved, as stated above.

§ 2.

Let now

$$\phi \equiv Ay_n^r + By_n^{r-1} + \cdots = 0$$

be such an equation of least order n, left invariant by a *primitive group*, which, according to definition, does not leave invariant a differential equation of the first order. It is supposed that ϕ does not contain any factor which is rational in $y_1, \cdots, y_n$. *We can show that* $n=2$.

The conditions for invariance are that $X^n\phi=0$ on account of $\phi=0$ for every infinitesimal transformation Xf of the group. Hence, as

$$X^n f \equiv (P_n y_n + Q_n)\frac{\partial f}{\partial y_n} + X^{n-1}f \qquad (n>1),$$

we have

$$AX^n\phi \equiv \phi(aAP_n + X^{n-1}A).$$

By taking the $(a-1)$-th partial derivative with respect to y_n of both members of these identities we get, after reduction,

$$AX^n(aAy_n + B) \equiv (aAy_n + B)(AP_n + X^{n-1}A). \tag{3}$$

The equation

$$aAy_n + B = 0$$

is therefore invariant. We may suppose that A and B have no factors in common containing derivatives.

Then follows from (3):

$$BX^{n-1}A \equiv A(aAQ_n + X^{n-1}B - BP_n). \tag{4}$$

Thus $A = 0$ is an invariant differential equation of less than the n-th order, contrary to hypothesis, unless A is a function of x, y only. Let us therefore put $B = aAC$, C being integral in $y_1, y_2, \cdots, y_{n-1}$. The invariant equation is therefore reducible to

$$y_n + C = 0,$$

and the equations (4) reduce to

$$-P_nC + Q_n + X^{n-1}C \equiv 0. \tag{5}$$

Suppose $n > 3$. Putting $C = ay'_{n-1} + \beta y'^{-1}_{n-1} + \cdots$, we find that, if $t > 1$,

$$(tP_{n-1} - P_n)a + X^{n-2}a = 0,$$

so that $a = 0$ would be an invariant differential equation, unless $a \equiv 0$. Indeed, a could not be a function of x, y merely, as not every

$$tP_{n-1} - P_n \equiv \xi_y y_1 + \xi_x - (t-1)(n\xi_y y_1 + n\xi_x - \eta_y)$$

could be free from y_1 for a primitive group, where not every ξ is a function of x alone.*

Thus,

$$C \equiv ay_{n-1} + \beta.$$

Substituting in (5) we find the conditions

$$(P_{n-1} - P_n)a - \frac{n(n+1)}{2}\xi_y y_2 + R_n(x, y, y_1) + X^{n-2}a \equiv 0. \tag{6}$$

By assuming a, which is integral in $y_1, y_2, \cdots, y_{n-2}$, to be of the form

$$\gamma y'_k + \delta y'^{-1}_k + \cdots,$$

* If every ξ were a function of x only, the group would leave invariant a differential equation of the first order, namely, $dx/dy = 0$.

y_k being the derivative of highest order contained in a, and substituting in (6), we find that $k = 2$ and $s = 1$, besides the conditions

$$\gamma(-P_n + P_{n-1} + P_2) - \frac{n(n+1)}{2}\xi_y + X^1\gamma = 0. \tag{7}$$

Let

$$\gamma = \epsilon y_1^r + \zeta y_1^{r-1} + \cdots.$$

Substituting in (7) and selecting the coefficients of y_1^{r+1} we get the equations

$$-(2+r)\xi_y\epsilon = 0,$$

which are impossibilities for every ξ_y that is not zero. It is to be noticed that ϵ could not be zero and therefore also γ, as then (7) would be impossible.

Thus $n > 3$ is not the least order of an invariant differential equation of the required form for a primitive group.

The case $n = 3$, when treated in the same manner, will give similar results. The coefficients of γ and ξ_y in the equations corresponding to (7) will here be slightly different from those in (7).

Thus $n = 2$. Equations (5) are here

$$-P_2C + Q_2 + X^1C \equiv 0,$$

from which we find

$$C \equiv ay_1^3 + \beta y_1^2 + \gamma y_1 + \delta,$$

a, β, γ, δ being functions of x, y only. The invariant differential equation is therefore

$$y_2 + ay_1^3 + \beta y_1^2 + \gamma y_1 + \delta = 0, \tag{8}$$

the reduction of which is our next object.

§ 3.

A three-parametric group, whose infinitesimal transformations are indicated by $Xf \equiv \xi p + \eta q$, leaves invariant a differential equation of the first order. For, if the determinant of the nine quantities ξ, η, η' is not zero identically, we have such an invariant equation as explained in the beginning of § 1, which will be of the first order in this case; and if the determinant vanishes identically, we have at least one differential invariant of the first order, J say, and therefore an invariant equation $f(J) = 0$.

A primitive group must therefore be at least four-parametric. Now, Dr. F. ENGEL has shown* that a group with more than three parameters contains at least one pair of independent infinitesimal transformations which are permutable.

* *Zur Theorie der Zusammensetzung der endlichen continuirlichen Transformationsgruppen*, Leipziger Berichte, 1886, pp. 83-94.

Such a pair, forming a two-parametric subgroup, can, by a proper choice of the variables, be given by

$$q, xq; \quad \text{or} \quad p, q.^*$$

In the first case it appears from a theorem by LIE† that the equation (8) can, by a transformation of the form

$$\bar{x} = x, \quad \bar{y} = y - \phi(x),$$

be changed into the following:

$$\bar{y}_2 = 0.$$

This transformation will change q, xq into $\bar{q}$, $\bar{x}\bar{q}$, where $\bar{q}$ stands for $\partial f/\partial \bar{y}$.

In the second case, by imposing upon the equation (8) the conditions that it should be left invariant by the infinitesimal transformations p, q, we find that the coefficients α, β, γ, δ are constants. Then, if $\alpha \neq 0$, by choosing for new variables

$$x = y - sx, \quad \bar{y} = x, \tag{9}$$

where

$$s^3\alpha + s^2\beta + s\gamma + \delta = 0,$$

we get a differential equation of the form

$$\bar{y}_2 + \alpha'\bar{y}_1^2 + \beta'\bar{y}_1 + \gamma' = 0 \tag{10}$$

in place of (8), α', β', γ' being constants.

The complete integral of this equation is readily found, and can in all cases be thrown into the form

$$\phi(\bar{x}, \bar{y}) + k_1\psi(\bar{x}, \bar{y}) + k_2 = 0,$$

k_1 and k_2 being the constants of integration. Then by changing to new variables x, y according to the relations

$$y = \phi(\bar{x}, \bar{y}), \quad x = \psi(\bar{x}, \bar{y}), \tag{11}$$

the equation (10) will reduce to

$$y_2 = 0.$$

Applying the combined transformations (9) and (11) to the subgroup p, q, we find this changed into one or other of the following subgroups, corresponding to the different forms of the functions ϕ and ψ of (11):

$$p + \alpha xq, q; \quad xp, yq; \quad p, yq; \quad xp, q;$$

α being a constant.

---

* *Differentialgleichungen mit bekannten infinitesimalen Transformationen*, p. 425.

† *Differentialgleichungen mit* etc., p. 429.

Hence, *by a proper transformation of the variables* x, y, *the primitive groups of the plane can be transformed into groups leaving invariant the differential equation*

$$\frac{d^2y}{dx^2} = 0,$$

and possessing one or other of the following subgroups:

$$q, xq; \quad q, p + axq; \quad xp, yq; \quad p, yq; \quad xp, q.$$

On account of the invariance of the equation $d^2y/dx^2 = 0$, *these transformed groups are projective.**

§ 4.

The problem of finding the primitive groups of the plane is thus reduced to one of finding all the different subgroups of the general projective group

$$p, \ q, \ yp, \ xq, \ xp, \ yq, \ x^2p + xyq, \ xyp + y^2q$$

which do not leave invariant a differential equation of the first order, and which contain the two-parametric subgroups given above.

Before we proceed it might be well to state a few well known propositions about the infinitesimal transformations of a group.

A set of r independent infinitesimal transformations generate an r-parameter group when, and only when, the commutator (Klammerausdruck) of any pair of these, as Xf and Yf:

$$(Xf, Yf) \equiv X(Yf) - Y(Xf),$$

is expressible as a sum of the given infinitesimal transformations multiplied by constants.

If Xf and Yf are any two infinitesimal transformations of a group, then is $aXf + bYf$ also an infinitesimal transformation of the group, a and b being constants.

A primitive group, which is projective, cannot leave invariant a given point in the plane, for otherwise the differential equation of the first order defining the ∞^1 straight lines through that point would be invariant.

Now, let the primitive group required contain the subgroup q, xq. The transformation

$$x = \frac{x'}{y'}, \quad y = \frac{1}{y'}$$

does not alter the form of the differential equation $d^2y/dx^2 = 0$ and it changes the

**Continuierliche Gruppen*, p. 35.

subgroup q, xq into the subgroup $x^2p + xyq, xyp + y^2q$, after we drop the accents.

One of the additional infinitesimal transformations of the group must be of the form

$$ap + bq + (cx + ey)p + (gx + hy)q \equiv Xf,$$

where a and b are not both zero, since it is necessary to have at least one infinitesimal transformation which does not leave invariant the point $(0, 0)$.

A linear transformation of the variables can now be chosen such that $ap + bq$ is changed into p and the subgroup $x^2p + xyq, xyp + y^2q$ is left unaltered, besides the differential equation $d^2y/dx^2 = 0$.

If now

$$Xf \equiv p + (cx + ey)p + (gx + hy)q,$$

we find by forming the following commutators

$$(Xf, x^2p + xyq) \equiv 2xp + yq + \alpha(x^2p + xyq) + \beta(xyp + y^2q) \equiv Yf,$$

$$(Xf, xyp + y^2q) \equiv yp + \alpha'(x^2p + xyq) + \beta'(xyp + y^2q),$$

$$\big((Xf, Yf) + Xf, Yf\big) \equiv 6p + 3\alpha(2xp + yq) + (4\beta + 2e)yp + \alpha''(x^2p + xyq) + \beta''(xyp + y^2q),$$

that the group required must contain the infinitesimal transformations

$$p, yp, 2xp + yq, x^2p + xyq, xyp + y^2q,$$

which will be found to generate a group. This group does not leave invariant a differential equation of the first order and is therefore primitive.

If the required group contains more infinitesimal transformations, we can assume these to be of the form $(a + bx + cy)q$. By building commutators as above we find that two cases present themselves, viz:

one additional infinitesimal transformation only, namely, yq;

three additional infinitesimal transformations, q, xq, yq.

We have now three types of primitive groups, which, by the transformation

$$x = \frac{x'}{y'}, \quad y = \frac{1}{y'},$$

become the following (dropping the accents):

$$\begin{aligned} &p, \ q, \ yp, \ xq, \ xp - yq; \\ &p, \ q, \ yp, \ xq, \ xp, \ yq; \\ &p, \ q, \ yp, \ xq, \ xp, \ yq, \ x^2p + xyq, \ xyp + y^2q. \end{aligned} \tag{12}$$

The two subgroups p, yq; xp, q are transformed one into the other by exchanging the variables. We need therefore consider only one of these, as p, yq.

The transformation $x = x'/y'$, $y = 1/y'$, used above, changes this subgroup into yp, $xp + yq$, after dropping the accents. As in the case considered above, we must have an infinitesimal transformation of the form

$$ap + bq + (cx + ey)q + g(x^2p + xyq) + h(xyp + y^2q) \equiv Xf,$$

where a and b are not both zero. Then from

$$\big(xp + yq, (xp + yq, Xf) - Xf\big) \equiv 2(ap + bq),$$
$$(yp, ap + bq) \equiv -bp,$$

we see that the required group must contain the infinitesimal transformation p, whether b be zero or not. We have now the subgroup p, yp which, by exchanging the variables, becomes the subgroup q, xq. This has been dealt with above.

Next, let the required group contain the subgroup xp, yq. We must have an infinitesimal transformation of the form

$$ap + bq + cyp + exq + g(x^2p + xyq) + h(xyp + y^2q) \equiv Xf,$$

a and b not being both zero.

From the commutators

$$\big(xp + yq, (xp + yq, Xf) - Xf\big) \equiv 2(ap + bq),$$
$$(xp, ap + bq) \equiv -ap, (yq, ap + bq) \equiv -bq,$$

we find that the group must contain one of the subgroups p, yq, or q, xp, which have just been considered.

Finally, let the subgroup $p + axq$, q belong to the group required. One of the additional infinitesimal transformations may be of the form

$$b(x^2p + xyq) + c(xyp + y^2q) + (gx + hy)p + (lx + ky)q \equiv Xf,$$

where b and c are not both zero.

Forming the commutators

$$\big(p + axq, (p + axq, Xf)\big) - (2b + ah)(p + axq)$$
$$\equiv -3a(b + ah)xq + 3acxp + aq \equiv Yf,$$
$$(p + axq, Yf) - 3ac(p + axq) \equiv -6a^2cxq + \beta q,$$
$$(q, Xf) - h(p + axq) \equiv (b - ah)xq + c(xp + 2yq) + lq,$$

we find that, if $a \neq 0$, we have the subgroup q, xq: and if $a = 0$, we see from

$$(-cp + bq, Xf) \equiv (bx + cy)(-cp + bq) + \gamma p + \delta q.$$

that we have the subgroup $-cp+bq$, $(bx+cy)(-cp+bq)$, which can be transformed into the subgroup q, xq by a linear transformation of the variables. But the primitive groups containing the subgroup q, xq have been determined above.

There remains to be determined the primitive groups whose infinitesimal transformations—at least four in number (§ 3)—are of the form

$$p+axq,\ q,\ (bx+cy)p+(ex+gy)q.$$

If we bear in mind that all the coefficients of p cannot be functions of x only, as we should then have an invariant differential equation of the first order $dx/dy=0$, and that all the coefficients of q cannot be functions of y only, as then the equation $dy/dx=0$ would be invariant, we find by building commutators that the first two groups of (12) will result. Accordingly:

All the primitive groups of the plane are represented by the types (12).*

*Cf. *Continuierliche Gruppen*, p. 360.

DETERMINATION OF ALL THE GROUPS OF ORDER p^m WHICH CONTAIN THE ABELIAN GROUP OF TYPE $(m-2, 1)$, p BEING ANY PRIME*

BY

G. A. MILLER

Introduction.

It is well known that there are just three abelian groups of order p^m which contain the group of type $(m-2, 1)$; viz. the groups of types $(m-1, 1)$, $(m-2, 2)$, $(m-2, 1, 1)$.† In what follows we shall therefore assume that the groups under consideration are non-abelian. The common abelian subgroup of type $(m-2, 1)$ will be represented by H. The operators of the required groups must transform those of H according to a subgroup of order p in the group of isomorphisms (I) of H. Our first problem is to determine all the operators of order p which are contained in I. Since the groups of order p^4 are well known we shall always assume $m > 4$. The groups which contain a cyclic invariant subgroup of order p^{m-2} are also known ‡, but we shall re-determine them in connection with the other groups since this will not materially affect the work.

The group H contains one cyclic characteristic subgroup of each order of the form p^a $(a = 1, 2, \cdots, m-3)$. It also contains one non-cyclic characteristic subgroup of each order of the form p^{a_1} $(a_1 = 2, 3, \cdots, m-2)$. Its other subgroups are cyclic and non-characteristic when p is odd. There are $p-1$ of these of each order of the form p^{a_2} $(a_2 = 2, 3, \cdots, m-3)$ while there are p such subgroups of each of the orders p and p^{m-2}. When $p = 2$ these subgroups of orders 2 and 2^{m-2} are the only non-characteristic subgroups of H. All the non-characteristic subgroups of the same order are conjugate under the holomorph of H. In fact, all the operators of the highest order in the non-characteristic subgroups of the same order are conjugate under this holomorph.

We proceed to prove several theorems applying to every abelian group (A), which will be employed in what follows. The operator of the group of isomor-

* Presented to the Society (Chicago) April 6, 1901. Received for publication March 13, 1901.

† FROBENIUS and STICKELBERGER, Crelle's Journal, vol. 84, 1879, p. 217.

‡ BURNSIDE, *Theory of groups of finite order*, 1897, p. 75.

phisms (A') of A, which transforms each operator of highest order in A into its β-th power must transform every operator of A into its β-th power since A is generated by its operators of highest order. Let s_1, s_2 be two operators of highest order (γ) in A and let $t^{-1}s_1t = s_1^\alpha$, $t^{-1}s_2t = s_2^\beta$ ($\beta \not\equiv a$, mod γ). The independent generators of the smallest group containing s_1 and s_2 may be represented by s_1 and s_3. From the equations just mentioned we have $t^{-1}s_3t = s_3^{\beta'}$ ($\beta' \not\equiv a$, mod the order of s_3). Since $t^{-1}s_1s_3t = s_1^\alpha s_3^{\beta'}$ is not a power of s_1s_3, t does not transform each subgroup of A into itself. Hence it follows that *the only operators of A' which transform each subgroup of A into itself are those which transform each operator of A into the same power.*

It will now be proved that the only invariant operators of A' are those which transform each operator of A into the same power of itself.* Let t' be any operator of A' which does not transform each operator of A into the same power. If t' permutes some of the subgroups of highest order in A it is not commutative† with the operator of A' which transforms into itself a generator of one of these subgroups (K') without transforming into itself a generator of the subgroup into which t' transforms K'. If t' does not permute any of the subgroups of highest order in A it must transform the generator of at least two of them into different powers and hence it is not commutative with the operator which transforms one of these two generators into the other. *The number of invariant operators of A' is therefore equal to the number of natural numbers which do not exceed the highest order of the operators of A and are prime to this order,* and the quotient group of A' with respect to these invariant operators is the group according to which the subgroups of A are transformed by A'.

§ 1. *Group of isomorphisms of H.*

Since each of the $p^{m-2}(p-1)$ operators of order p^{m-2} in H can be made to correspond with every other operator of this order‡ and any holomorphism of H is completely determined by such a correspondence and a correspondence between two of the $p(p-1)$ operators of order p which are not found in the characteristic subgroup of order p, the order of I is $p^{m-1}(p-1)^2$. By raising all the operators of H to each one of the $p^{m-3}(p-1)$ powers which are prime to p and not greater than p^{m-2} we obtain the holomorphisms of H which correspond to the cyclic subgroup of I all of whose operators are invariant under I.§ To determine the remaining operators of I we employ a method which

*Cf. Transactions of the American Mathematical Society, vol. 1, 1900, p. 396.

†Cf. Bocher, Mathematische Annalen, vol. 49, 1897, p. 134.

‡Transactions of the American Mathematical Society, vol. 1, 1900, p. 397.

§Transactions of the American Mathematical Society, vol. 1, 1900, p. 396. When $p=2$ this group of order 2^{m-3} is of type $(m-4, 1)$. In what follows p is assumed to be odd unless the contrary is stated.

was explained in a recent number of the Bulletin of the American Mathematical Society.*

The group I contains an invariant subgroup (I_2) of order p^2 which is composed of all the operators of I which are commutative with each operator of the characteristic subgroup (in H) of order p^{m-2}. These isomorphisms can be obtained by making H isomorphic with its subgroups of order p, the characteristic subgroup of order p^{m-2} corresponding to identity. Hence I_2 is of type $(1, 1)$. By making H isomorphic with its characteristic subgroup of order p with respect to one of its cyclic subgroups of order p^{m-2} we observe that I contains operators of order p which are not contained in I_2. This proves that I contains a subgroup (I_3) of order p^3 which includes no operator of order p^2. We shall now prove that I_3 is the non-abelian group of order p^3 which contains no operator of order p^2.

Let P_1, P_2 represent two independent generators of H, the order of P_1 being p^{m-2}, and let $P_1^{p^{m-3}} = P_3$. The first four of the following equations result directly from the given holomorphisms of H while the last two show that two operators (t_1, t_2) of I_3 are not commutative and hence that I_3 is non-abelian.

$$t_1^{-1}P_2t_1 = P_2, \quad t_1^{-1}P_1t_1 = P_2P_1, \quad t_2^{-1}P_1t_2 = P_1, \quad t_2^{-1}P_2t_1 = P_3P_2,$$

$$t_2^{-1}t_1^{-1}P_1t_1t_2 = P_3P_2P_1, \quad t_1^{-1}t_2^{-1}P_1t_2t_1 = P_2P_1.$$

It has been observed that I contains a cyclic group of order $p^{m-3}(p-1)$ all of whose operators are invariant. Each operator of this group transforms every subgroup of H into itself. The corresponding quotient group of I is of order $p^2(p-1)$. We proceed to prove that this group can be represented as an intransitive substitution group of degree $2p$, formed by establishing a (p, p) correspondence between two metacyclic groups of degree p. Each of the two sets of p non-characteristic subgroups of the same order in H is transformed by I according to the metacyclic group of degree p. As cyclic groups of order p^{m-2} are transformed into each other by operators of I which transform each one of the cyclic groups of order p into itself, the subgroup of H must be transformed according to the said intransitive substitution group. Since the order of this group is the order of the quotient group of I with respect to its operators which transform each subgroup of H into itself, this intransitive group is the quotient group in question. As this quotient group contains only one subgroup of order p^2, I contains only one subgroup (K) of order p^{m-1}.

The subgroup K contains no invariant operators besides those of the given cyclic subgroup of order p^{m-3} since it must include I_3. The group of cogredient isomorphisms of K is therefore the same as that of I_3. Hence it contains the abelian subgroup of type $(m-3, 1)$ which includes I_2. The remaining operators of K are the products of this abelian subgroup and the first $p-1$ powers

* Vol. 6, 1900, p. 337.

of any operator (t_4) of I_3 which is not included in I_2. Let t_5 be any operator of the group of type $(m-3, 1)$ first mentioned. From the properties of the operators of I_3 it follows that

$$(t_5t_4)^p = t_5t_4t_5t_4\cdots p \text{ times} = t_5t_4t_5t_4^{-1}t_4^2t_5t_4^{-2}\cdots t_4^{-1}t_5t_4 = t_5^p uu^2u^3\cdots u^{p-1},$$

u being an invariant operator of I_3 which may be identity. Since t_5t_4 is of the same order as t_5 whenever t_5 is not identity I *contains no operator of order* p *besides those of* I_3.

As it is necessary to consider all the operators of I_3 in the construction of the groups under consideration we proceed to determine its sets of operators which are conjugate under I. It has been observed that $p-1$ of its operators are invariant under I and that its subgroup I_2 is also invariant under I. It is clear that I_3 contains p other subgroups of order p^2. It will be proved that one of these is always invariant under I while the other $p-1$ are conjugate when $p > 3$. From the quotient group of I with respect to its invariant operators and from the properties of I_3 it follows that each of the p^3-p non-invariant operators of I belongs to a set of $p(p-1)$ conjugates under I. Hence all the non-invariant operators of the invariant subgroups of order p^2 are conjugate.

We proceed to consider the holomorphisms obtained by making H isomorphic with its non-cyclic subgroup of order p^2, the cyclic characteristic subgroup of order p^{m-3} corresponding to identity and the non-cyclic subgroup of order p^{m-2} corresponding to the group generated by P_3. From the following equations it may be observed that each one of these $p(p-1)^2$ holomorphisms gives rise to an operator of order p in I. In these equations t_6 is the operator of I which corresponds to the holomorphism under consideration.

$$t_6^{-1}P_1t_6 = P_2P_1, \quad t_6^{-2}P_1t_6^2 = P_3P_2^2P_1, \quad t_6^{-3}P_1t_6^3 = P_3^3P_2^3P_1,$$

$$t_6^{-u}P_1t_6^u = P_3^{1+2+\cdots+u-1}P_2^uP_1, \; t_6^{-p}P_1t_6^p = P_1, \; t_6^{-1}P_2t_6 = P_3P_2, \; t_6^{-p}P_2t_6^p = P_3^pP_2.$$

We have now found all the holomorphisms of H which correspond to the operators of I_3. In the p^2 holomorphisms which correspond to I_2 each of the operators of the non-cyclic subgroup of order p^{m-2} corresponds to itself. In $p(p-1)$ others each operator of a cyclic group of order p^{m-2} corresponds to itself while in $p(p-1)^2$ just found the only operators which correspond to themselves are those of the cyclic characteristic subgroup of order p^{m-3}. The $p(p-1)$ operators, each of which is commutative with some operator of order p^{m-2}, together with the invariant operators of I_3 must correspond to an invariant subgroup in the quotient group of I with respect to its invariant operators. Hence *they form an invariant subgroup of order* p^2.

We shall now prove that the $p(p-1)$ subgroups, generated by the $p(p-1)^2$ operators of I_3 which transform only p^{m-3} operators of H into themselves, are

conjugate under I whenever $p > 3$. Let t_7 be any operator of I which satisfies the two conditions

$$t_7^{-1} P_1 t_7 = P_1, \quad t_7^{-1} P_2 t_7 = P_3^{\alpha} P_2^{\beta}, \qquad \beta \not\equiv 0, \mod p.$$

The transform of t_6 with respect to t_7 satisfies the following condition:

$$(t_7^{-1} t_6 t_7)^{-1} P_1 t_7^{-1} t_6 t^7 = P_3^{\alpha} P_2^{\beta} P_1.$$

If we assign to t_7 all its values subject to the given conditions we obtain the $p(p-1)$ holomorphisms which correspond to the conjugates of t_6. It remains to prove that the one which corresponds to t_6^2 is not in this set. If it were we should have $P_3^{\alpha} P_2^{\beta} = P_3 P_2^2$ and hence $\alpha = 1$, $\beta = 2$. This would lead to the following equations which are impossible when $p > 3$:

$$(t_7^{-1} t_6 t_7)^{-1} P_2^2 t_7^{-1} t_6 t_7 = P_3 P_2^2 = P_3^4 P_2^2 = t_6^{-2} P_2^2 t_6^2.$$

Since I does not transform t_6 into t_6^2, its quotient group with respect to its invariant operators can be obtained by establishing such a (p, p) correspondence between two metacyclic groups of degree p as leads to a group having only two invariant subgroups of order p.* This is clearly impossible when $p = 3$. In this special case each of the subgroups of order p^2 in I_3 is therefore invariant under I while, in all other cases, I_3 contains only two invariant subgroups of order p^2 and a set of $p-1$ conjugate subgroups of this order.

The rest of this section will be devoted to the study of I when $p = 2$.

We may assume $m > 5$ since all the groups of order 32 are known. Now I is of order 2^{m-1} and it is isomorphic to the abelian group of type $(m-5, 1)$ with respect to its operators which are commutative with each operator of the cyclic subgroup of order 2^{m-3} in H which is generated by P_1. Hence I_2 and I_3 may be found in the same manner as when p is odd, the latter being the group of order 8 containing five operators of order 2 and two of order 4 since it is non-abelian and contains I_2.

It has been observed that the 2^{m-3} invariant operators of I form a group of type $(m-4, 1)$. Since this cannot contain more than 2 operators of I_3, I and I_3 must have the same group of cogredient isomorphisms. Hence I contains an abelian group of type $(m-4, 1, 1)$ including I_2, and the rest of its operators transform just half of the operators of this group into themselves multiplied by the invariant operator (s_4) of I_3. Since s_4 is commutative with all the operators

* By establishing a (p, p) correspondence between two metacyclic groups of degree p, written in distinct sets of letters, we obtain as many intransitive groups of order $p^2(p-1)$ as there are operators in the group of isomorphisms of the cyclic group of order $p-1$ diminished by one-half of the number of its operators whose orders exceed two. Cf. American Journal of Mathematics, vol. 21, 1899, pp. 292 and 293. In one of these groups each subgroup of order p is invariant while each of the others contains only two invariant subgroups of this order. The latter exists only whenever $p > 3$.

of H whose order is less than 2^{m-2} it is the square of the operators of order four in the given group of type $(m-4, 1, 1)$. We shall soon observe that all the operators of order 4 in I have the same square. Let s_5 be any operator of the mentioned group of type $(m-4, 1, 1)$ and suppose that s_5 is not invariant in I. Also let s_6 represent an operator of order 2 in I_3 that is not also contained in I_2. From the equation $(s_5 s_6)^2 = s_5 s_6 s_5 s_6 = s_4 s_5^2$ it follows that we obtain additional operators of order two when s_5 is of order four and only then. When s_5 is of order two we obtain the additional operators of order four. Since s_4 is the square of all of these, *all the operators of order four in I have the same square.*

From the results of the preceding paragraph it follows that all the operators of order two in I are contained in a subgroup (I_5) of order 32 which includes all the operators of order four in I. Also I_5 contains 16 operators of order four and 15 of order two. As the latter are not all commutative to each other they must generate I_5. Moreover it is clear that I_5 is the direct product of its operator of order two which transforms each operator of H into its inverse and the group (I_4) of order 16 which contains just 7 operators of order two that generate it.* Since I transforms the four non-characteristic subgroups of H according to the intransitive substitution group of degree and order four, each of its non-invariant operators has just two conjugates. Hence I_5 *contains three invariant operators of order two and six pairs of conjugates.*

§ 2. *Outline of the method employed to construct the groups in question.*

If a group of order p^m is represented as a regular substitution group, every subgroup of order p^{m-1} contains p systems of intransitivity. We may therefore suppose that each of the groups under consideration contains the intransitive substitution group formed by writing H in the regular form in p distinct sets of letters and establishing a simple isomorphism between these p regular groups $H_1, H_2, \cdots, H_p$. In what follows H will represent this intransitive group of type $(m-2, 1)$ while t will be used to represent the substitution of order p which merely permutes the corresponding letters of $H_1, H_2, \cdots, H_p$. Hence t is commutative with every substitution of H. Also, I will represent the intransitive substitution group which is formed by writing the maximal subgroup of degree $p^{m-1} - a$ $(a > 0)$ of the holomorph of H_1† in the p distinct sets of letters involved in H; so that t is also commutative with each substitution of I. Finally, i will be used to represent some substitution of order p that is contained in I, while I_2, I_3, I_4, and I_5 will represent the same subgroups of I as in the preceding section.

*Quarterly Journal of Mathematics, vol. 28, 1896, p. 271, no. 9.

†Bulletin of the American Mathematical Society, vol. 5, 1899, p. 245.

Now H and ti clearly generate a group (G) of order p^m. All the other groups of this order which transform H in the same manner as G does, may be generated by H and sti, where s is commutative with each substitution of H; for, if such a group is generated by H and r, where r permutes the systems of H in the same way as t does, then $rt^{-1}i^{-1} = s$ must be commutative with each substitution of H, i. e., $r = sti$. Hence s is a substitution of the direct product of $H_1, H_2, \cdots, H_p$. We may assume that s involves no letters except those of H_1; for if it involved the letters of $H_\alpha (p \geqq \alpha > 1)$ but not those of $H_{\alpha+\beta} (\beta > 0)$, the transform of the group generated by H and sti with respect to the component of s which involved the letters of H_α only would be a conjugate group in which s would be replaced by a substitution involving the letters of $H_{\alpha-1}$ but not those of $H_{\alpha-1-\beta}$. Since the other conditions named would not be affected it is proved that *we may assume that s involves the letters of H_1 only.* On this account it will hereafter be denoted by s_1 and its conjugates under the group generated by t will be denoted by $s_1, s_2, \cdots, s_p$.

By hypothesis $(s_1 ti)^p$ is some substitution of H, but

$$(s_1 ti)^p = s_1 ti \cdot s_1 ti \cdots = s_1 i s_p i s_{p-1} i \cdots s_2 i = s_1 s_p' s_{p-1}' \cdots s_2',$$

where s_α' is the transform of s_α with respect to some power of i. From this it follows that s_1 *and i are always commutative.* Let r_1 represent a substitution of H that is transformed by i^{-1} into itself multiplied by a substitution u_1 of order p which is commutative to i. The conjugates of r_1 and u_1 with respect to the powers of t will be represented by $r_1, r_2, \cdots, r_p$; $u_1, u_2, \cdots, u_p$ respectively. It is easy to verify that

$$(r_1^{p-1} r_3 r_4^2 \cdots r_p^{p-2})^{-1} t i r_1^{p-1} r_3 r_4^2 \cdots r_p^{p-2} = r_1^{1-p} r_3^{-1} r_4^{-2} \cdots r_p^{2-p} r_2 r_3^2 \cdots r_p^{p-1} u_2 u_3^2 \cdots u_p^{p-1} ti$$

$$= r_1^{1-p} r_2 r_3 \cdots r_p u_2 u_3^2 \cdots u_p^{p-1} ti.$$

If we transform the last substitution successively by $u_p^{p-1}, u_{p-1}^{2p-3}, \cdots, u_3^{\frac{(p-1)(p-2)}{2}}$ and multiply the result by $r_1^{-1}, r_2^{-1}, \cdots, r_p^{-1}$, which is in H, we observe that the group generated by H and ti is conjugate to the one which is generated by H and $r_1^{-p} ti$; i. e., *no new group is obtained when s_1 is the pth power of a substitution of H_1 which is transformed by i into itself multiplied by a substitution of order p that is commutative with i.*

By interchanging systems of intransitivity of H, t can be transformed into any power while each of the substitutions i and s_1 is transformed into itself. Hence the $p-1$ powers of i which are prime to the order of i give rise to the same groups. It is also clear that each of the conjugates of i under I leads to the same groups. Moreover, if each of two groups generated by ti_α and ti_β respectively, contains only one subgroup that is simply isomorphic with H and if i_α is not conjugate with some power of i_β under I, then must the two groups

be distinct; for if they were simply isomorphic they could be transformed by operators of I so that the identical operators of H would correspond. The remaining operators of the group would have to transform H in the same manner; i. e., the group generated by i_α would be conjugate to the one generated by i_β. The main results of this section may be summarized as follows:

1. Only one operator from each set of subgroups of order p which are conjugate under I is to be used for i.

2. s_1 is a substitution of H_1 which is commutative with i.

3. s_1 is not the pth power of a substitution of H_1 which is either commutative with i or is transformed by i into itself multiplied by a substitution of order p that is commutative with i.

4. All the powers of s_1 which are prime to its order lead to the same group since it can be transformed into these powers by operators that are commutative with t and i.

§ 3. *Determination of the groups when p is odd.*

It was observed in § 1 that i is contained in the non-abelian group I_3 of order p^3 which contains no operator of order p^2. We shall begin with the case when i is one (i_1) of the $p-1$ operators of the invariant subgroup of order p. Hence i_1 is commutative with every operator of the non-cyclic subgroup of order p^{m-2} in H. From the equations

$$(P_1 t i_1)^p = P_1 t i_1 \cdot P_1 t i_1 \cdots p \text{ times} = P_1 i_1 \cdot P_1 i_1 \cdots t^p$$

$$= P_1 i_1 \cdot P_1 i_1^{-1} i_1^2 P_1 i_1^{-2} i_1^3 \cdots i_1^{p-1} P_1 i_1^{1-p} i_1^p = P_3 P_3^2 \cdots P_3^{p-1} P_1^p = P_1^p,$$

it follows that the product of $t i_1$ and any operator of H besides identity is of the same order as the operator of H. The group (G_1) generated by H and $t i_1$ has therefore the same number of operators of each order as the abelian group of type $(m-2, 1, 1)$. Two groups which have the same number of operators of each order may be called conformal.* Now G_1 contains just $p+1$ abelian subgroups of order p^{m-1}, p of these are of type $(m-2, 1)$ while the remaining one is of type $(m-3, 1, 1)$. As each of the former is transformed in the same manner by all the operators of G_1, we shall not arrive at G_1 when we use in place of i_1 an operator from another set of conjugates in I_3.

From § 2 it follows that $s_1 t i_1$ and H generate G_1 whenever s_1 is an operator of the characteristic cyclic subgroup of order p^{m-3} in H_1. If s_1 is the constituent of P_2 which involves the letters of H_1 only we clearly obtain a group (G_2) which contains no operator of order p besides those of H. Then G_2 contains the same number of operators of each order $> p^2$ as G_1 does. It also includes the same

*Bulletin of the American Mathematical Society, vol. 2, 1896, p. 140.

number of abelian subgroups of type $(m-2, 1)$ as G_1 does, but the group of type $(m-3, 1, 1)$ in G_1 is replaced by one of type $(m-3, 2)$ in G_2. These are the only abelian subgroups of order p^{m-1} that are contained in G_2 and all those of type $(m-2, 1)$ are transformed in the same way by all the operators of G_2. Since we obtain G_2 if we replace s_1 by any one of its $p-1$ powers multiplied into any operator of the cyclic characteristic subgroup of order p^{m-3} in H_1, i_1 gives rise to only the two groups G_1 and G_2.

We shall now consider the possible groups when i is one (i_2) of the $p(p-1)$ conjugate operators of I_2. i_2 and i_1 are commutative with the same operators of H. Just as in the preceding case we find that i_2 leads to two groups (G_3 and G_4) which are conformal with G_1 and G_2 respectively. In G_1 and G_2 all the cyclic subgroups of order p^{m-2} are invariant but in G_3 and G_4 these subgroups are conjugate in sets of p. Each of the groups G_3 and G_4 contains $p+1$ abelian subgroups of order p^{m-1}, p of these are of type $(m-2, 1)$ and are transformed by the remaining operators in the same way as H is transformed. Hence we cannot obtain any of the four groups just found by using for i an operator of I_3 which is not contained in I_2. By using for i one (i_3) of the $p(p-1)$ conjugate operators of the other invariant subgroup of I_3, when $p>3$, we clearly obtain only one additional group (G_5) that involves no operator whose order exceeds p^{m-2}. The group G_5 contains $p+1$ abelian subgroups of type $(m-2, 1)$ and it transforms each of its cyclic subgroups of order p^{m-2} into itself. It is the only group involving H that has invariant operators of order p^{m-2} without containing any operator of order p^{m-1}. It is clear that i_3 leads also to the non-abelian group (G_6) of order p^m * which involves an operator of order p^{m-1} and to no other group.

Each of the remaining $p(p-1)^2$ operators of I_3 transforms only p^{m-3} operators of H into themselves and the $p(p-1)$ subgroups which they generate are conjugate when $p>3$. Hence it is only necessary to consider one (i_4) of them when $p>3$. The group (G_7) generated by ti_4 and H contains only one abelian subgroup of order p^{m-1}. From the equation

$$(P_1 t i_4)^p = (P i_4)^p = P_1 i_4 P_1 i_4^{-1} i_4^2 P_1 i_4^{-2} i_4^3 \cdots i_4^{p-1} P_1 i_4^{1-p} = P_1^p \qquad \text{(when } p>3\text{)}$$

$$\text{or} = P_3 P_1^p \qquad \text{(when } p=3\text{)},$$

it follows that G_7 is conformal with G_1, G_3 and G_5. The following equations show that i_4 gives rise to only one group. In these equations r_1 represents the constituent of P_1 that is contained in H_1, while v_1 and w_1 represent the corresponding constituents of P_2 and P_3 respectively. The higher subscripts indicate the conjugates of these substitutions with respect to powers of t.

* BURNSIDE, *Theory of groups of finite order*, 1897, p. 76.

$$(r_1^{p-1}r_3r_4^2\cdots r_p^{p-2})^{-1}ti_4r_1^{p-1}r_3r_4^2\cdots r_p^{p-2}=r_1^{1-p}r_3^{-1}r_4^{-2}\cdots r_p^{2-p}i_4r_2r_3^2\cdots r_p^{p-1}t$$

$$=r_1^{1-p}r_2r_3\cdots r_pv_2v_3^2\cdots v_p^{p-1}ti=r_1^{1-p}r_2r_3\cdots r_pw_1^{-1-3-6\cdots\frac{-(p-1)(p-2)}{2}}ti_4$$

which, when multiplied by $r^{-1}r_2^{-1}r_3^{-1}\cdots r_p^{-1}$ of H, is equal to $r_1^{-p}ti$ if $p>3$, and to $w_1^{-1}r_1^{-p}ti$ if $p=3$.

When $p=3$, I_3 contains another invariant subgroup of order p^2. If i_5 represents one of its conjugate operators we obtain an additional group G_8 by multiplying H and ti_5. It may be proved in exactly the same manner as in the preceding case that i_5 leads to only one group, and it is, moreover, clear that G_8 and G_7 are conformal. Since each contains only one abelian subgroup of order p^{m-1} and since i_4 and i_5 are not conjugate under I, G_8 and G_7 cannot be conjugate.

§ 4. *Determination of the groups when p is even.*

In § 1 it was observed that all the operators of order two in I are contained in a group of order 32 which has three invariant operators (i_1, i_2, i_3) of order 2; viz., those which transform each operator of H into its $2^{m-3}+1$, $2^{m-3}-1$, $2^{m-2}-1$ powers respectively. The group (G_1) generated by H and ti_1 is conformal with the abelian group of type $(m-2, 1, 1)$ and each of its subgroups, with exception of the four of order two which are not in H, is invariant. The group G_1 contains three abelian subgroups of order p^{m-1}, two being of type $(m-2, 1)$ and the third of type $(m-3, 1, 1)$, and it transforms all the operators of the former into their $2^{m-3}+1$ power. In exactly the same manner as in the preceding section we obtain G_2 by multiplying H and s_1ti_1, s_1 being the constituent of P_2 which is in H_1. The group G_2 contains only three operators of order 2 and each of these is the square of four operators of order 4. The operators of its two abelian subgroups of type $(m-2, 1)$ are transformed in the same way as those of the corresponding subgroups of G_1.

The group (G_3) generated by H and ti_2 contains 2^{m-2} operators of each of the orders two and four in addition to the operators of H. All the operators of order four in G_3 have the same square. Let s_1 be the constituent of P_1 which involves the same letters as H_1. From the equation $s_1^{-1}ti_2s_1=s_1^{-1}s_2^{2^{m-3}-1}ti_2$ it follows that transforms of G_3 include H and $s_1^{2^{m-3}}ti_2$. Hence there is only one additional group (G_4) which transforms H in the same way as i_2 does, viz., the group generated by H and s_1ti_2, where s_1 represents the constituent of P_2 which is in H_1. All the operators of G_4 which are not found in H are of order 4 and the squares of these operators of order 4 are the two non-characteristic operators of order two in H, each of the latter being the square of one-half of the former.

The group H and ti_3 generate a group (G_5) composed of H and 2^{m-1} operators of order 2 which are conjugate in sets of 2^{m-3}. In fact, it is clear that in all

the groups to which i_2 and i_3 lead, all the operators not contained in H are conjugate in sets of 2^{m-3}. The group (G_6) generated by H and $s_1 t i_3$, where s_1 is the component of P_3 which is found in H_1, contains $2^{m-1}+4$ operators of order four. All of these have the same square. If s_1 is the component of P_2 which belongs to H_1, the group (G_7) is conformal with G_6 but its operators of order four which are not found in H have P_2 for their square. As the other operator of H_1 which is commutative with i_3 is conjugate under I with the last s_1, the seven groups just found are all those which depend upon i_1, i_2, i_4. In each of these seven groups every cyclic subgroup whose order exceeds four is invariant. It remains to find the groups which depend upon operators that are not invariant under I.

Let i_4 be one of the two non-invariant operators of I_3 which are commutative with each operator of the non-cyclic subgroup of order 2^{m-2} in H. Since the product of ti_4 into any operator of H is of the same order as this operator of H, the group (G_8) generated by H and ti_4 is conformal with G_1. G_8 contains three abelian subgroups of types $(2^{m-2}, 1)$ and $(2^{m-3}, 1, 1)$, there being two of the former type. The four cyclic subgroups of order 2^{m-2} form two pairs of conjugates. If s_1 represents the component of P_1 which is contained in H_1 we have $s_1^{-1} t i_4 s_1 = s_1^{-1} s_2 s_2' t i_4$, s_2' being the constituent of P_2 which is in H_2. Hence a transform of G_8 contains H and $s_1' s_1^{-2} t i_4$. We therefore obtain only one additional group (G_9) by multiplying H and $s_1 t i_4$, where s_1 is a substitution of H_1. It may be convenient to assume that G_9 is generated by H and $s_1 t i_4$, where s_1 is the constituent of P_2 which is in H_1. From this we observe that G_9 is conformal with G_2. Like G_8 it contains two pairs of conjugate cyclic subgroups of order 2^{m-2} and three abelian subgroups of order 2^{m-1}. One of these is, however, of type $(2^{m-3}, 2)$ instead of type $(2^{m-3}, 1, 1)$ in G_8.

Each of the other two conjugate operators of order two in I_3 is commutative with an operator of order 2^{m-2} in H. If i_5 represents one of these operators it is clear that H and ti_5 generate a group (G_{10}) which is conformal with G_1. Its subgroup of order 16 which includes all the operators of orders two and four is generated by its operators of order two. Its four cyclic subgroups of order 2^{m-2} are invariant, one being generated by an invariant operator while the operators of the others are transformed into their $2^{m-3}+1$ power. The other group (G_{11}) which involves invariant operators of order 2^{m-2} is generated by H and $s_1 t i_5$, where i_5 is the constituent of P_1 which is in H_1. Hence G_{11} contains operators of order 2^{m-1} which it transforms into their $2^{m-2}+1$ power. It includes just two cyclic subgroups of order 2^{m-1}.

The group (G_{12}) generated by H and ti_6, where i_6 is one of the two conjugate operators of order two in I which are commutative with each operator of the cyclic subgroup of order 2^{m-3} generated by $P_1^2 P_2$ without being commutative with any other operator of H, is conformal with G_1. Like G_3, G_4, G_5, G_6,

and G_7 this group and all those which follow contain only one abelian subgroup of order 2^{m-1} while each of the other six groups (G_1, G_2, G_8, G_9, G_{10}, G_{11}) contains just three abelian groups of this order. From the equation $s_1^{-1}ti_6s_1 = s_2's_1^{-1}s_2ti_6$, s_1 being an operator of order 2^{m-2} in H_1, it follows that i_6 leads to one other group of order 2^m. The four cyclic subgroups of order 2^{m-2} are conjugate in pairs in G_{12}.

We have now considered all the operators of order two which are contained in I_4 as well as the two additional operators i_3 and $i_1i_3 = i_2$. It remains to consider the three operators i_3i_4, i_3i_5, and i_3i_6. We shall represent these respectively by i_7, i_8, and i_9. The group H and ti_7 generate a group (G_{13}) which contains 2^{m-2} operators of each of the orders two and four in addition to H. H and s_1ti_7, where s_1 is the constituent of P_3 which is in H_1, generate a group G_{14} which contains only operators of order four in addition to H. Neither of these two groups contains an invariant cyclic subgroup of order 2^{m-2}. From the equations $s_1^{-1}ti_7s_1 = s_1^{-1}s_2^{-1}s_2'ti_7$, s_2' being the constituent of P_2 in H_2, it follows that i_7 gives rise to no additional group.

The group (G_{15}) generated by H and ti_8 is conformal with G_{13} but all its operators of order four have the same square while in G_{13} they have two distinct squares. Moreover, its invariant operators form a cyclic group of order four while they form a non-cyclic group of this order in G_{13}. The group H and s_1ti_8, where s_1 is the H_1 component of the invariant operator of order four in G_{15}, generate a group (G_{16}) which contains only operators of order eight in addition to those of H. The two cyclic subgroups of order 2^{m-2} are clearly invariant in each one of these two groups. That these two are the only groups to which i_8 gives rise follows from the equation $s_1^{-1}ti_8s_1 = s_1^{-1}s_2^{-1+2^{m-3}}$, s_1s_2 being an operator of order 2^{m-2} in H which is not commutative with i_5.

H and ti_9 generate a group (G_{17}) which contains, besides H, 2^{m-2} operators of order eight and 2^{m-3} operators of each of the orders two and four. Like the preceding four groups it contains just four invariant operators. That this is the only group to which i_9 leads follows from the equations $s_1^{-1}ti_9s_1 = s_1^{-1}s_2^{-1}s_2'$, where s_2' is of order four if s_1 is of order 2^{m-2}; and of order one when s_1 is of order 2^{m-3} and not commutative with i_6. Hence the number of the groups of order 2^m which contain the abelian group of type $(m-2, 1)$ is twenty, three being abelian and seventeen being non-abelian. Eleven of the latter involve invariant cyclic subgroups of order 2^{m-2}.

Summary.

There are three abelian groups of order p^m that contain a subgroup of type $(m-2, 1)$, viz., those of types $(m-1, 1)$, $(m-2, 2)$, $(m-2, 1, 1)$. These exist for every prime p.

When $p > 3$ there are seven other groups. The number of their operators of each order is exhibited by the following table:

Groups	Operators				
	Order	p	p^2	p^α	p^{m-1}
G_1, G_3, G_5, G_7	Number	p^3-1	$p^3(p-1)$	$p^{\alpha+1}(p-1)$	0
G_2, G_4	"	0	p^4-1	$p^{\alpha+1}(p-1)$	0
G_6	"	p^2-1	$p^2(p-1)$	$p^\alpha(p-1)$	$p^{m-1}(p-1)$

$(p>3) \quad (2<\alpha<m-1)$.

The first four groups may be distinguished by the following properties: G_7 contains only one abelian subgroup of order p^{m-1} while each of the others contains three such subgroups; G_5 is the only one which contains invariant operators of order p^{m-2}; G_1 contains cyclic invariant subgroups of order p^{m-2} while G_3 does not; G_2 and G_4 can be distinguished by the same property as was employed to distinguish G_1 and G_3. When $p=3$ there is one more group, which is conformal with G_6. It is not simply isomorphic with G_6 since these two groups contain only one abelian subgroup of order p^{m-1} which they transform according to non-conjugate subgroups of I.

The following table exhibits the orders of the operators of the 17 non-abelian groups of order 2^m which include an abelian group of type $(m-2, 1)$:

Groups	Operators					
	Order	2	4	8	2^α	2^{m-1}
G_1, G_8, G_{10}, G_{12}	Number	7	8	16	$2^{\alpha+1}$	0
G_2, G_9	"	3	12	16	$2^{\alpha+1}$	0
G_3, G_{13}, G_{15}	"	$2^{m-2}+3$	$2^{m-2}+4$	8	2^α	0
G_4, G_6, G_7, G_{14}	"	3	$2^{m-1}+4$	8	2^α	0
G_5	"	$2^{m-1}+3$	4	8	2^α	0
G_{11}	"	3	4	8	2^α	2^{m-1}
G_{17}	"	$2^{m-3}+3$	$2^{m-3}+4$	$2^{m-2}+8$	2^α	0
G_{16}	"	3	4	$2^{m-1}+8$	2^α	0

$(3<\alpha<m-1)$.

The conformal groups may be distinguished by the following properties: G_{10} is the only one of the first four which contains invariant operators of order 2^{m-2}; G_{12} contains only one abelian group of order 2^{m-1} while G_1 and G_8 contain three such groups; in G_1 all the cyclic subgroups of order 2^{m-2} are invariant while they are conjugate in sets of two in G_8; G_2 and G_9 may be distinguished in the same way as G_1 and G_8; G_{13} contains two operators of order two which are squares of operators of order four while in G_3 and G_{15} all the

operators of order four have the same square; in G_3 the invariant operators form a non-cyclic subgroup of order four while they form a cyclic subgroup of this order in G_{15}; in G_{14} the two cyclic subgroups of order 2^{m-2} are conjugate while they are invariant in G_4, G_6, and G_7; in G_4 the operators of order 4 have three distinct squares, in G_6 they have the same square, and in G_7 they have two distinct squares.

CORNELL UNIVERSITY, *March*, 1901.

ON A FUNDAMENTAL PROPERTY OF A MINIMUM IN THE CALCULUS OF VARIATIONS AND THE PROOF OF A THEOREM OF WEIERSTRASS'S*

BY

W. F. OSGOOD

The notion of the minimum of an integral in the calculus of variations is analogous to that of the minimum of a function of a real variable. The function, $f(x)$, is thought of as continuous and it is said to have a minimum at the point $x = a$ when, for all points of the neighborhood of a with the one exception of a itself, the value of $f(x)$ is greater than $f(a)$; i. e., when the relation holds:

$$f(x) > f(a), \quad 0 < |x - a| \leqq h.$$

This is, however, only a partial description of the behavior of the function in the neighborhood of the point a, as is shown by a simple example of a discontinuous function for which the above relation is true. Let the function $\phi(x)$ be defined as follows:

$\phi(x) = 1$, when x is irrational;

$\phi(x) = 1/q$, when $x = \pm p/q$, where p/q is a positive fraction in its lowest terms (in particular, an integer).

$\phi(0) = 0$.

Then $\phi(x)$ has a minimum at the point $x = 0$, for

$$\phi(x) > \phi(0), \quad x \neq 0;$$

and yet, in the neighborhood of an arbitrarily chosen value of x, the lower limit of $\phi(x)$ is 0.

A continuous function $f(x)$ has, then, a further property in the neighborhood of a point $x = a$ at which it is a minimum, which may be expressed as follows:

Let δ be any positive quantity less than h. Then there exists a positive quantity ϵ such that the lower limit of $f(x)$ for values of x in the above neighborhood of a for which $|x - a| > \delta$ is at least as great as $f(a) + \epsilon$; or

$$f(x) \geqq f(a) + \epsilon, \quad \delta < |x - a| \leqq h.$$

* Presented to the Society April 27, 1901. Received for publication May 6, 1901.

In the calculus of variations the integral

$$I = \int_{x_0}^{x_1} F(x, y, y')\, dx$$

or, in the parametric form,

$$I = \int_{t_0}^{t_1} F(x, y, x', y')\, dt$$

is said to be made a minimum by the functions $x = \phi(t)$, $y = \psi(t)$ corresponding to a curve C if its value J for C is less than its value I for any other curve C of a specified class (K) of curves which all lie in the neighborhood T of C (Fig. 1):

$$I > J.$$

WEIERSTRASS* has given a sufficient condition that this may be the case. The example of the function $\phi(x)$ above considered raises, however, the question: *Assuming that Weierstrass's sufficient condition is fulfilled, so that no curve of the class (K) distinct from C will give the integral I so small a value as J, may we not still have a set of curves belonging to this class, $\bar{C}_1, \bar{C}_2, \cdots$, which do not cluster about C as their limit and which have the property that, if $I_1, I_2, \cdots$ denote respectively the values of the integral I formed for these curves,*

$$\lim_{n=\infty} I_n = J\ ?$$

In particular, it is conceivable that these curves might approach a curve C of the class (K) as their limit, distinct from C. For this curve, the value of I would, of course, not be $\lim_{n=\infty} I_n = J$.

This question I answer in the negative, showing in § 1 by means of the theorem of § 2 that, if $\mathfrak{T}$ denotes an arbitrarily small neighborhood of C lying wholly within T, the lower limit of I formed for all the curves of the class (K) that do not lie wholly within $\mathfrak{T}$ is greater than J by a positive quantity ϵ, so that

$$I \geqq J + \epsilon,$$

when $\bar{C}$ does not lie wholly in $\mathfrak{T}$.

By means of these results it is possible to deal with a problem of importance concerning the scope of the class of curves (K). What curves shall be admitted to this class? The larger the class (K), the more general the property of the integral which is made a minimum by C; for the integral is thus compared

* In his lectures. For a statement of the condition in the parametric case, see § 1 below. This is in substance KNESER'S form of the sufficient condition: KNESER, *Lehrbuch der Variationsrechnung*, chap. 3. Other forms of the sufficient condition, in particular, HILBERT'S form, have recently been treated in a paper by the writer: Annals of Mathematics, 2nd series, vol. 2, no. 3 (1901), p. 105.

with the curves of a more extended class. Hence it is natural to seek to enlarge the class (K) until limitations are reached which are dictated by the *nature of the problem*, not by the *exigencies of the proof*. For example, consider the integral

$$I = \int_{t_0}^{t_1} \sqrt{x'^2 + y'^2}\, dt.$$

This integral represents the length of the curve C when C has a continuous tangent. But the notion of length is not restricted to curves of this class. In fact, a curve may have a length and yet fail to have a tangent in a set of points everywhere dense along the curve. For, the length of a curve may be defined as the upper limit of the lengths of the inscribed polygons, in case a finite upper limit exists, i. e., as

$$L \sum \sqrt{\Delta x_i^2 + \Delta y_i^2},$$

where $\Delta x_i = \phi(\tau_{i+1}) - \phi(\tau_i)$, $\Delta y_i = \psi(\tau_{i+1}) - \psi(\tau_i)$, and the vertices $\tau_0 = t_0$, $\tau_1, \cdots \tau_n = t_1$ are chosen in all possible ways.* JORDAN has shown that the necessary and sufficient condition that such an upper limit exist is that the functions ϕ, ψ both belong to the class of functions with limited variation (fonctions à variation bornée).†

The above generalization of the integral that represents the length of a curve with continuous tangent suggests a corresponding generalization of the general integral

$$I = \int_{t_0}^{t_1} F(x, y, x', y')\, dt.$$

Let C be a continuous curve lying in T and connecting the points A and B. Since C will not in general have a tangent, the proper integral I will not in general have a meaning. Inscribe a polygon in C, as in the case of the arc integral, and, remembering that the function F is homogeneous of the first degree in its last two arguments, form the sum

* This is in substance SCHEEFFER's definition. The form, which is a particularly elegant one is due to PEANO: *Applicazioni geometriche del calcolo infinitesimale*, 1887, p. 161; *Lezioni di analisi infinitesimale*, vol. 1, 1893, p. 171; Annali di matematica, 2nd ser., vol. 23 (1895), p. 153. It is analogous to PEANO's definition of the upper definite integral; cf. *Lezioni* (just cited), § 103; Annali, loc. cit. See also HILBERT's definition at the close of this paper.

† JORDAN, *Cours d'analyse*, vol. 1, 2d ed., 1893, § 67. An example of a continuous curve, $y = \Phi(x)$, which has a length, but which does not have a tangent for any rational value of the abscissa can easily be given. Let $f(x) = x \sin \log x$, $x \neq 0$; $f(0) = 0$, and let $\varphi(x) = f(\sin \pi x)$. Then the function $\Phi(x)$, given by the series

$$\Phi(x) = a_1\varphi(x) + a_2\varphi(2!\,x) + a_3\varphi(3!\,x) + \cdots,$$

where the coefficients a_n are suitably chosen ($a_n = 1/n!^2$ will suffice), is a continuous function of limited variation. It is readily seen that this function, however, fails to have a derivative when x is rational. — The graph of WEIERSTRASS's continuous function that has no derivative is an example of a continuous curve that does not have a length.

$$\sum F\left(x_i, y_i, \frac{\Delta x_i}{\Delta t_i}, \frac{\Delta y_i}{\Delta t_i}\right)\Delta t_i = \sum F(x_i, y_i, \Delta x_i, \Delta y_i).$$

If this sum converges toward one and the same limit when $n = \infty$ no matter how the polygons be chosen, provided merely that their sides all converge toward 0, then this curve shall be included in the class (K) and the value of the limit of this sum shall be taken as the generalized integral I. If C has a continuous tangent, this value is the same as that of the proper integral I. Thus we have a natural limitation for the class (K) of curves C to be admitted to consideration. WEIERSTRASS, to whom the generalization here considered is due, states in his lectures the theorem that when his sufficient condition for a minimum is fulfilled, the integral will still be a minimum, even when its definition is thus generalized and its value for C is compared with its value for any curve of this class (K). I have never seen a proof of this theorem. In § 3 of the present paper a proof is given.

The theorem of § 1 is proven only for the simplest integral of the calculus of variations; but it is true for other integrals, for example, for the double integral

$$\int\int F(x, y, z, p, q)\, dx\, dy,$$

and it probably holds in general. In the case of a weak minimum, the neighborhood T of C must be replaced by the *engere Nachbarschaft* * of C.

In the proof which HILBERT has recently given of DIRICHLET'S Principle † a continuous function z of the variables x, y is defined as $\lim_{n=\infty} F_n(x, y)$ and it has to be shown that z satisfies LAPLACE'S equation at all interior points of the regi . The theorem of § 1, stated for double integrals, may be employed to estal ish this point. A corresponding step has to be taken in the proof of the existence of a shortest line on a surface, and the theorems of this paper are applicable in this case, too.

§ 1. A CHECK FOR THE VALUE OF THE INTEGRAL I EXTENDED ALONG A CURVE $\bar{C}$ NOT LYING WHOLLY WITHIN THE NEIGHBORHOOD $\mathfrak{T}$ OF THE EXTREMAL C.

1. *The quantity H, expressed in terms of the curvilinear coördinates of the field.* — Consider the integral

$$I = \int_{t_0}^{t_1} F(x, y, x', y')\, dt$$

* ZERMELO, *Untersuchungen zur Variationsrechnung*, Berlin dissertation, 1894.

† Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 8 (1899) p. 184.

extended along the curve

$$C: \quad x = \phi(t), \quad y = \psi(t), \quad t_0 \leqq t \leqq t_1,$$

that connects the fixed points A, with the coördinates (x_0, y_0), $(t = t_0)$, and B, with the coördinates (x_1, y_1), $(t = t_1)$. Here, ϕ, ψ shall be continuous functions of t for the values in question, and shall have continuous first derivatives that do not vanish simultaneously for any of the above values of t. The function F shall, together with all its partial derivatives of the first and those of the second order that present themselves, be a single valued continuous function of the four arguments (x, y, x', y') regarded as independent variables, where (x, y) is any point of a certain region of the (x, y)-plane, within which the curve C must lie, and x', y' have any values except the one pair $(0, 0)$; furthermore, since the integral is to depend only on the curve C, not on the particular choice of the parameter t, the function F shall be homogeneous of the first degree in the last two arguments. Finally, we assume that F is positive for all values of the arguments considered.

A necessary condition for a minimum is that the curve C satisfy the differential equations

$$F_x - \frac{d}{dt} F_{x'} = 0, \quad F_y - \frac{d}{dt} F_{y'} = 0, \qquad \left(F_x = \frac{\partial F}{\partial x}, \text{ etc.}\right).$$

We assume that this condition is fulfilled by the extremal* C, and that C is a curve which does not cut itself between A and B; furthermore, that a field exists surrounding C. More precisely, we assume that a one-parameter family of curves exists:

$$x = \xi(t, a), \quad y = \eta(t, a),$$

where the functions ξ, η have the following properties:

(a) These functions, together with their first partial derivatives and the cross derivatives ξ_{ta}, η_{ta}, are continuous functions of the independent varibles (t, a) at all points of the region

$$R: \qquad \bar{t}_0 \leqq t \leqq \bar{t}_1, \quad |a - a_0| \leqq \kappa,$$

where $\bar{t}_0 < t_0$, $t_1 < \bar{t}_1$, and $\kappa > 0$; $\bar{t}_0$, $\bar{t}_1$, κ are constants. When $a = a_0$, the equations represent a curve without multiple points, which passes through A and B; and when a has any constant value, the corresponding curve is an extremal. In particular, when $a = a_0$, the curve is the extremal C.

* Any curve that satisfies these equations is denoted by KNESER as an extremal; *Lehrbuch der Variationsrechnung*, p. 24.

(*b*) The determinant

$$\Delta = \begin{vmatrix} \xi_t & \xi_a \\ \eta_t & \eta_a \end{vmatrix} \neq 0,$$

no matter for what point of R it be formed.

Let κ be chosen small enough so that all the curves will be free from multiple points. Two curves corresponding to distinct values of a not only will not intersect each other, but, if the two values of a differ from each other by an infinitesimal a, the distance $PP' = \beta$ from a point P of the one curve to the point of intersection P' of the normal at P with the other curve will be an infinitesimal that is uniformly of the first order; in other words,

$$ma \leqq \beta \leqq Ma,$$

where m, M are positive constants, i. e., independent of a, t, a.

The proof of these statements follows the same lines as the proof in a similar case considered in the writer's paper on *Sufficient Conditions in the Calculus of Variations*.* KNESER† has given a proof by means of power series on the assumption that all the functions considered are analytic. This assumption is unfortunate, inasmuch as it introduces restrictions without being accompanied by any simplification.

By means of the family of extremals $a = \text{const.}$, a system of curvilinear coördinates of the field is introduced (KNESER, loc. cit.), $v = a$ being taken as one of the coördinates and

$$u = \int_{t'}^{t} F\big(\xi(t, a), \eta(t, a), \xi_t(t, a), \eta_t(t, a)\big)\, dt$$

being chosen as the other. The lower limit of integration, t', is so taken that the point $x = \xi(t', a)$, $y = \eta(t', a)$ lies on the curve through A which cuts all the extremals $a = \text{const.}$ transversely. In terms of the new coördinates let

$$F(x, y, dx, dy) = G(u, v, du, dv),$$

so that

$$I = \int_{t_0}^{t_1} F(x, y, x', y')\, dt = \int_{t_0}^{t_1} G(u, v, u', v')\, dt,$$

where I is taken along the curve C. The quantity, H, the study of which is the main object of this paragraph, is then given by the formula

$$H = I - J = \int_{t_0}^{t_1} [G(u, v, u', v') - u']\, dt.$$

*Annals of Mathematics, 2nd series, vol. 2, No. 3 (April, 1901), p. 105.

†loc. cit., chap. 3.

This quantity will be positive, and hence the integral I will be made a minimum by C, when the WEIERSTRASSIAN invariant F_1 (KNESER, p. 52) is always positive in R, and this we assume to be the case. This completes the sufficient condition for a minimum.

2. *The determination of a lower limit (Abschätzung) for the value of H for the simplest class of curves $\bar{C}$.* — Consider first the case that $u' > 0$ throughout the whole interval $t_0 \leqq t \leqq t_1$. Then u may be chosen as the parameter t and, setting

$$G(u, v, 1, s) = g(u, v, s),$$

we have

$$[G(u, v, u', v') - u']\, dt = [g(u, v, s) - 1]\, du.$$

KNESER shows that

$$g(u, v, 0) = 1, \quad g_s(u, v, 0) = 0,$$

so that

$$H = \tfrac{1}{2} \int_{u_0}^{u_1} g_{ss}(u, v, \vartheta s) s^2\, du,$$

the coördinates of A and B being respectively (u_0, v_0) and (u_1, v_0). The function $g_{ss}(u, v, s)$ is a continuous function of its three arguments, when s has any value whatever and (u, v) is any point of the region

$$T: \qquad \bar{u}_0 \leqq u \leqq \bar{u}_1, \qquad |v - v_0| \leqq \bar{\kappa},$$

where the constants $\bar{u}_0$, $\bar{u}_1$, $\bar{\kappa}$ are so chosen that to each point of T corresponds a point of R, while the arc AB of C lies wholly within T. This function is positive for all such values of the arguments, since the invariant F_1 is always positive. It may, however, have 0 for its lower limit when $|s| = \infty$. Let the minimum value of this function, when (u, v) is an arbitrary point of T and $|s| \leqq \mu$ (μ being a constant to be determined later) be denoted by M.* If we impose on the curve $\bar{C}$ the further restriction that s shall, at no one of its points, exceed μ in absolute value, then we have

$$H \geqq M \int_{u_0}^{u_1} s^2\, du.$$

Now mark off an arbitrary neighborhood of the arc AB of C lying within the field T and consisting of the region

$$\mathfrak{T}: \quad u_0 - L \leqq u \leqq u_1 + L, \quad |v - v_0| \leqq L; \quad L < 3\sqrt{2}(u_1 - u_0).$$

* It will be necessary to consider later the minimum value of the function $\frac{1}{2}\bar{g}_{ss}(u, v, s)$ where $G(u, v, -1, s) = \bar{g}(u, v, s)$, for the same values of the arguments, and so we will choose M once for all as the smaller of these two minimum values.

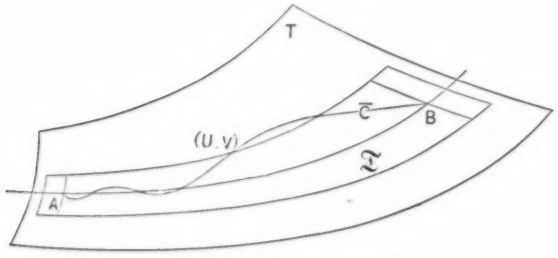

FIG. 1.

If the curve C, restricted as above, does not lie wholly within $\mathfrak{T}$ and (U, V) is a point of C lying on the boundary of $\mathfrak{T}$, so that $|V| = L$, then, by the theorem of § 2

$$\int_{u_0}^{U} s^2\,du \geqq \frac{L^3}{54(u_1 - u_0)^2},$$

and since

$$\int_{u_0}^{u_1} s^2\,du > \int_{u_0}^{U} s^2\,du,$$

it follows that

$$(A_1) \qquad H > \frac{ML^3}{54(u_1 - u_0)^2}.$$

3. *Continuation; the general case.* — We have thus obtained the desired inequality, or approximate evaluation, (Abschätzung) for H for the simplest class of curves C. The most general curve

$$C: \qquad x = \phi(t), \quad y = \psi(t); \qquad t_0 \leqq t \leqq t_1,$$

connecting A and B, having at each point (extremities included) a continuous tangent, and lying in T, but not wholly in $\mathfrak{T}$, can be broken up into a finite number of pieces of the following three classes:

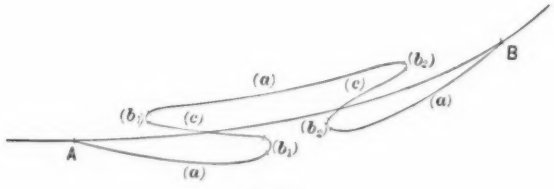

FIG. 2.

$$(a) \qquad u' > 0, \quad |s| \leqq \mu;$$

$$(b) \qquad \begin{cases} (b_1) & v' > 0 \text{ and, when } u' = 0, \ |s| > \tfrac{1}{2}\mu; \\ (b_2) & v' < 0 \text{ and, when } u' = 0, \ |s| > \tfrac{1}{2}\mu; \end{cases}$$

$$(c) \qquad u' < 0, \quad |s| \leqq \mu.$$

Instead of $\frac{1}{2}\mu$, any other positive quantity less than μ would answer. — The integral H can then be decomposed in a corresponding manner into the three sums:

$$(1) \qquad H = \Big(\sum_{(a)} + \sum_{(b)} + \sum_{(c)}\Big)\int [G(u, v, u', v') - u']dt.$$

We proceed to obtain approximate evaluations for these sums, and hence for H.

Consider a term of the first sum

$$H_{a_i} = \int_{t_i}^{t_{i+1}} [G(u, v, u', v') - u']\, dt.$$

Here, $u = \phi(t)$, $u' > 0$ and hence the variable of integration, t, can be replaced by u. We will introduce, however, not u but w, where

$$w = u + a_i$$

and a_i is a constant to be determined later. We have then

$$(2) \qquad H_{a_i} = \int_{w_i}^{w_{i+1}} \tfrac{1}{2} g_{ss}(u, v, s)\, s^2\, dw \ \geqq \ M \int_{w_i}^{w_{i+1}} s^2\, dw.$$

Next consider a term of the third sum

$$H_{c_k} = \int_{t_k}^{t_{k+1}} [G(u, v, u', v') - u']dt$$

and replace t by w, where

$$w = -u + a_k$$

and a_k is a constant to be determined later. Let

$$G(u, v, -1, s) = \bar{g}(u, v, s).$$

Then, as before,

$$\bar{g}(u, v, 0) = 1, \quad \bar{g}_s(u, v, 0) = 0.$$

Hence

$$(3) \quad H_{c_k} = \int_{w_k}^{w_{k+1}} [\tfrac{1}{2}\bar{g}_{ss}(u, v, \vartheta s)\, s^2 + 2]dw \ \geqq \ M \int_{w_k}^{w_{k+1}} s^2\, dw + 2(w_{k+1} - w_k).$$

Finally, consider a term of the second sum for which $v' > 0$. Replace t by v, and denote the smaller of the minimum values of the two functions

$$G(u, v, r, 1), \quad G(u, v, r, -1),$$

when (u, v) is any point of T and $r = du/dv (= s^{-1}$ when $u' \neq 0$ and $v' \neq 0)$, $|r| \leqq 1$, by N. Then, since the maximum value of $|r|$ is $2/\mu$, if μ is so chosen (1) that $2/\mu \leqq 1$ and (2) that $2/\mu < N$, we obtain for H_{b_j} the following relation.

$$H_{b_j} = \int_{t_j}^{t_{j+1}} [G(u, v, u', v') - u'] dt = \int_{v_j}^{v_{j+1}} [G(u, v, r, 1) - r] dv$$

(4)

$$> \left(N - \frac{2}{\mu}\right)(v_{j+1} - v_j).$$

The same formula also holds for a term for which $v' < 0$, since in that case the left hand member is positive, the right hand member negative. Furthermore, the same reasoning shows that, when $v' < 0$, the quantity $v_{j+1} - v_j$ then becoming negative,

$$(4_1) \qquad H_{b_j} > -\left(N - \frac{2}{\mu}\right)(v_{j+1} - v_j).$$

We are now ready to obtain the final approximate evaluation of the quantity H. Assume that the curve $\bar{C}$ leaves the strip $\mathfrak{T}$, and let (U, V) be a point of $\bar{C}$ on the boundary of $\mathfrak{T}$. Two cases arise:

$$(A) \qquad u_0 \leqq U \leqq u_1, \quad |v| = L;$$

$$(B) \qquad U > u_1 \quad \text{or} \quad U < u_0.$$

Case (A). We begin by disposing of a simple special case, namely, that in which

$$\sum_k 2(w_{k+1} - w_k),$$

the summation applying to those arcs of $\bar{C}$ which belong to class (c), exceeds an arbitrarily preassigned quantity, which we take for simplicity as L. Observe that the equation of the boundary of $\mathfrak{T}$ for which u is a maximum is $u = u_1 + L$. Here, we have at once that

$$(A_2) \qquad H > L.$$

This case being disposed of, we assume that

$$\sum_k 2(w_{k+1} - w_k) \leqq L$$

and that the point (U, V) is the boundary of an interval of some one of the three classes. This condition can be fulfilled by cutting an interval in two, if necessary; both pieces belonging, then, to the same class. We now determine the constants a_i, a_k in such a manner that those intervals

$$w \leqq w \leqq w_{i+1}, \quad w_k \leqq w \leqq w_{k+1}$$

which correspond to the arcs of C lying between A and (U, V) will follow each other without overlapping, their ends just touching; the initial value of w in the first interval shall be chosen as w_0. Denote the final value of w in the last interval by $\mathfrak{w}$. We have, then, from (2) and (3), suppressing whatever terms of the form $2(w_{k+1} - w_k)$ may arise on the right hand side of (3), — this only strengthens the inequality: —

$$\left(\sum_{(a)} + \sum_{(c)}\right)\int [G(u, v, u', v') - u']dt > M\int_{w_0}^{\mathfrak{w}} s^2 dw. \tag{5}$$

The function s that here enters is a single valued function of w continuous at all points of the interval $w_0 \leqq w \leqq \mathfrak{w}$ with the exception of a finite number of points; and s approaches a limit when w approaches any one of these exceptional points from one side only. Let

$$f(w) = \int_{w_0}^{w} s\, dw.$$

Then $v = f(w)$ is a continuous curve made up (1) of the pieces of class (a) carried in the field of the curvilinear coördinates (u, v) to a new part of the field; (2) of the pieces of class (c), first reflected in a line $u =$ constant (or $v =$ constant) and then carried like the former pieces to a new part of the field. Let $f(\mathfrak{w})$ be denoted by W; $f(w_0) = 0$. We have the relation:

$$W + \sum(v_{j+1} - v_j) = V. \tag{6}$$

Collecting and combining results we infer from (4), (4_1), (5) that†

$$H > M\int_{w_0}^{\mathfrak{w}} s^2 dw + \left(N - \frac{2}{\mu}\right)|x|, \tag{7}$$

where

$$x = \sum(v_{j+1} - v_j).$$

Relation (7) we now proceed to transform by means of the theorem of § 2, which enables us to write the inequality

$$\int_{w_0}^{\mathfrak{w}} s^2 dw \;\geqq\; \frac{|W|^3}{54(u_1 + L - u_0)^2},$$

provided that $|W| < 3\sqrt{2}(u_1 + L - u_0)$; a condition that will surely be satisfied when $|W| \leqq L$. We will assume this to be the case, since otherwise we could choose $0 < \mathfrak{w}_1 < \mathfrak{w}$ so that $|f(\mathfrak{w}_1)| = L$, and then we should have

* We may think of the field T as mapped on a rectangle by means of the equations $u = x$, $v = y$. The curves in question thus go over into curves lying in this rectangle and the transformation is then an ordinary euclidean rigid motion. The corresponding actual transformations in the field T now become apparent to the intuition. They are given by the formulas:

(1) $u = w - a_i$, $v = v$; (2) $u = -w - a_k$, $v = v$.

† In obtaining this relation *either* (4) *or* (4_1) is to be used for all values of j.

$$\int^{w} s^2\,dw > \int_{u_0}^{w_1} s^2\,dw \;\geqq\; \frac{L^3}{54(u_1+L-u_0)^2},$$

whence relation (A_3) would follow at once.

We have then, since $W = V - x$,

$$H > \frac{M|V-x|^3}{54(u_1+L-u_0)^2} + \left(N - \frac{2}{\mu}\right)|x|, \tag{8}$$

or, writing

$$\frac{M}{54(u_1+L-u_0)^2} = A, \quad \left(N-\frac{2}{\mu}\right) = B,$$

$$H > A|V-x|^3 + B|x|. \tag{8'}$$

A and B are both positive quantities.

When V and x have opposite signs, the inequality (8′) will only be strengthened by replacing x by 0. Since $|V| = L$,

$$H > AL^3. \tag{A_3}$$

When V and x have the same sign, we can replace (8′) by the inequality

$$H > A(L-\xi)^3 + B\xi, \tag{9_1}$$

when $|x| = \xi < L$; and by the inequality

$$H > A(\xi - L)^3 + B\xi, \tag{9_2}$$

when $\xi \geqq L$. To get the desired approximate evaluation, replace ξ in each of these formulas by a value that will make the right hand side of the inequality a minimum. The curve

$$y = A(L-\xi)^3 + B\xi$$

has its minimum value in the interval $0 \leqq \xi < L$ at the point

$$\xi = L - \sqrt{B/3A},$$

provided that this is a point of the interval, i. e., provided $L \geqq \sqrt{B/3A}$; otherwise, at the point $\xi = 0$. In the first case the minimum value is

$$B[L - \tfrac{2}{3}\sqrt{B/3A}];$$

in the second case it is AL^3.

When $\xi \geqq L$, the minimum is given by putting $\xi = L$, and is BL.

Collecting results, we have

$$(A_4) \quad \begin{cases} H > B(L - \frac{2}{3}\sqrt{B/3A}), & \text{when } 0 \leqq \xi < L, \quad \sqrt{B/3A} \leqq L; \\ H > AL^3, & \text{when } 0 \leqq \xi < L, \quad \sqrt{B/3A} > L; \\ H > BL, & \text{when } \xi \geqq L. \end{cases}$$

It remains to consider *Case* (B). It is sufficient to assume that $U > u_1$ for the case that $U < u_0$ can be dealt with in a similar manner, the points A and B merely interchanging their rôles. The curve C in this case crosses the curve $u = u_1$ (let $t = \tau_1$, be the value of t corresponding to such a point) and meets one of the three lines

$$v = L, \qquad u_1 < u \leqq u_1 + L;$$

$$v = -L, \quad u_1 < u \leqq u_1 + L;$$

$$v_0 - L \leqq v \leqq v_0 + L, \quad u = u_1 + L,$$

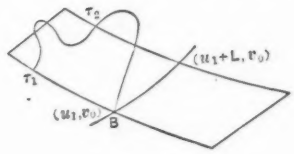

FIG. 3.

in the point (U, V) (let $t = \tau_2$ be the value of t corresponding to such a point). Let the minimum distance from the point B to the above three lines be denoted by D. Then there exists a positive constant, k, such that

$$D > kL,$$

no matter what value L may have subject to the conditions of the problem we are considering. *

We may write H in the form:

$$H = \int_{t_0}^{\tau_1} [G(u, v, u', v') - u']\, dt + \int_{\tau_1}^{\tau_2} G(u, v, u', v')\, dt$$

$$+ \int_{\tau_2}^{t_1} G(u, v, u', v')\, dt.$$

The first of these integrals is positive or zero, since it is the difference between the integral I taken along a curve $\bar{C}$ drawn in the field from A to the curve $u = u_1$, which cuts the extremals of the field transversely, and the integral I taken from A to B along C. The second integral is also positive since G is positive. Discard these first two integrals and then introduce in the third as parameter the arc s of C. We have then

$$(A_5) \qquad H > \int_{\sigma_2}^{\sigma_1} G(u, v, \cos\phi, \sin\phi)\, ds > k\Gamma L,$$

* For the region about A a corresponding pair of quantities D', k' must be considered, and so we will choose D, k at once each as the smaller one of the two quantities to be considered.

where $\Gamma > 0$ denotes the minimum value of $G(u, v, \cos\phi, \sin\phi)$ when (u, v) is any point of T and ϕ has any value whatever.

This completes the approximate evaluation. The final result may be stated as follows:

If L is any positive quantity so chosen that no point of the region $\mathfrak{T}$ will be exterior to the region T and that $L < 3\sqrt{2}(u_1 - u_0)$, and if E denotes the smallest of the quantities entering on the right hand sides of the inequalities for H: (A_1), (A_2), (A_3), (A_4), (A_5); then the value of the integral I extended along any curve C whatsoever that does not lie wholly within $\mathfrak{T}$ exceeds the value of this integral extended along C by a quantity greater than E; or

$$H > E$$

for such a curve C.

The curves C considered in the foregoing investigation were assumed to have a continuous tangent throughout their whole course inclusive of their extremities. It is easy to see that the above results still hold for curves C made up of a finite number of pieces, each of which enjoys the above properties. In particular, C may be made up of a finite number of pieces of straight lines.

§ 2. The Fundamental Theorem.

4. Fundamental Theorem. *Let $f(x)$ be a single valued continuous function of x in the interval $a \leqq x \leqq b$, and let $f(x)$ have a continuous derivative $f'(x)$ at all points of this interval. Let*

$$|f(l) - f(a)| = L > 0, \qquad a < l \leqq b,$$

Then

$$\int_a^l f'(x)^2\,dx \geqq \frac{L^3}{54(b-a)^2},$$

provided that

$$L < 3\sqrt{2}(b-a).$$

*The theorem still holds under the more general hypothesis that $f(x)$ is continuous throughout the interval $a \leqq x \leqq b$ and that this interval can be broken up into a finite number of subintervals throughout each of which, inclusive of its extremities, $f(x)$ has a continuous derivative.**

We may restrict ourselves to the case that $f(l) - f(a) > 0$, since the case that $f(l) - f(a) < 0$ can be referred to the former case by setting $f(x) = -\phi(x)$. We may furthermore assume that $f'(x) \geqq 0$, since, if $f'(x) < 0$ in parts of the interval, we may define a function $F(x)$ by the conditions:

* Since this theorem merely serves the purpose of a lemma for the proof of the main theorems of this paper, no attempt is made to state the theorem in more general form than is needed for present purposes.

$$F(x) = f(x) \text{ for values of } x \text{ for which } f'(x) \geqq 0:$$

$$F(x) = 0 \quad \text{for values of } x \text{ for which } f'(x) < 0,$$

and this function will have the property that

$$F(l) - F(a) = L > L.$$

If l' be so chosen that $F(l') = L$ and $0 < l' < l$, then

$$\int_a^l f'(x)^2\,dx > \int_a^l F'(x)^2\,dx > \int_a^{l'} F'(x)^2\,dx \geqq \frac{L^3}{54(b-a)^2}.$$

Thus the theorem will hold in all cases if it holds when $f(l) > f(a)$ and $f'(x) \geqq 0$, $a \leqq x \leqq b$.

Finally, we shall for simplicity confine ourselves to the proof of the theorem under the hypothesis that $f'(x)$ is continuous throughout the whole interval, for the more general case presents no principal difficulty.

5. LEMMA. *If $f'(x)$ is a single-valued, continuous function of x throughout the interval $\alpha \leqq x \leqq \beta$, and if*

$$0 \leqq f'(x) \leqq 1$$

at all points of the interval; if, furthermore,

$$f(\beta) - f(\alpha) = \lambda > 0 \quad \text{and if} \quad \lambda < \beta - \alpha,$$

then

$$\int_\alpha^\beta f'(x)^2\,dx > \tfrac{1}{8}\frac{\lambda^3}{(\beta-\alpha)^2}.$$

Let γ, τ be defined by the equations:

$$\beta - \alpha = \tau,$$

$$\int_\alpha^\beta f'(x)\,dx = f(\beta) - f(\alpha) = \gamma(\beta - \alpha),$$

so that $\gamma\tau = \lambda$, $0 < \gamma < 1$.

Let the points of the interval (α, β) be divided into two sets:

(s_1), the points in which $f'(x) \geqq \frac{1}{2}\gamma$, $\qquad \kappa'$

(s_2), the points in which $f'(x) < \frac{1}{2}\gamma$; $\qquad \tau - \kappa'$

and let κ', $\tau - \kappa'$ denote respectively the internal content of (s_1) and the external content of (s_2).* The quantity κ' varies with different functions $f'(x)$,

* Cf. JORDAN, *Cours d'analyse*, vol. 1, 2d ed., 1893, p. 28, § 36.

and we wish to determine its lower limit. Consider the function $\phi(x)$ defined as follows:

$$\phi(x) = 1, \quad a \leqq x \leqq a + \kappa; \qquad \phi(x) = \tfrac{1}{2}\gamma, \quad a + \kappa < x \leqq \beta,$$

where κ is so chosen that

$$(1) \qquad \begin{cases} \int_a^\beta \phi(x)dx = \lambda, \quad \text{i. e.,} \quad \kappa + \tfrac{1}{2}\gamma(\tau - \kappa) = \lambda, \\ \kappa = \dfrac{\lambda}{2 - \gamma} > \tfrac{1}{2}\lambda. \end{cases}$$

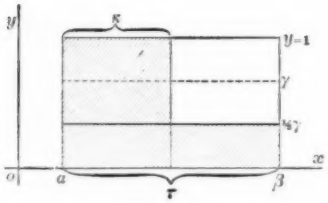

FIG. 4.

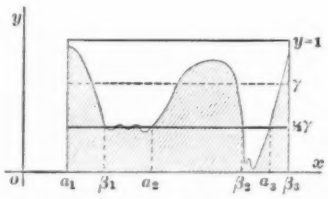

FIG. 5.

This function ϕ is, because of its discontinuity at $x = a + \kappa$, an impossible function $f'(x)$; but there exist functions $f'(x)$ that differ from $\phi(x)$ but slightly, and it seems likely that the number κ thus defined is the lower limit of all the κ''s. This we will now show to be the case.

A positive number ϵ having been chosen at pleasure, it is possible to divide the interval (a, β) into a finite number of subintervals (a_1, β_1), (a_2, β_2), $\cdots$, (a_n, β_n), consisting exclusively of points of (s_1) and such that

$$\kappa' - \epsilon < \sum_{i=1}^{n} (\beta_i - a_i) < \kappa'.$$

The sum of the lengths of the remaining intervals then satisfies the relation:

$$\tau - \kappa' < \sum_{i=0}^{n} (a_{i+1} - \beta_i) < \tau - \kappa' + \epsilon,$$

where $\beta_0 = a$ and $a_{n+1} = \beta$. In particular, it may happen, as in the case indicated in the figure, that $a_1 = a$, so that the first interval, $(a_1 - \beta_0)$, may be suppressed. A similar remark applies to the last interval when $\beta_n = \beta$. Furthermore, the above intervals may be so chosen that, δ being a second arbitrary positive number,

$$f'(x) < \tfrac{1}{2}\gamma + \delta,$$

at all points of each of the intervals (β_i, a_{i+1}). This follows from the uniform continuity of the function $f'(x)$.

From the formula

$$\int_\alpha^\beta f'(x)\,dx = \sum_{i=1}^{n}\int_{\alpha_i}^{\beta_i} f'(x)\,dx + \sum_{i=0}^{n}\int_{\beta_i}^{\alpha_{i+1}} f'(x)\,dx,$$

we now infer, since $f'(x) \leqq 1$ in the interval (α_i, β_i) and $f'(x) < \frac{1}{2}\gamma + \delta$ in the interval (a_{i+1}, β_i), that

$$\lambda < \kappa' + (\tfrac{1}{2}\gamma + \delta)(\tau - \kappa' + \epsilon). \tag{2}$$

From (1) and (2) it follows that

$$\kappa < \kappa' + (\epsilon, \delta),$$

where (ϵ, δ) is a quantity that can be made arbitrarily small by a suitable choice of ϵ and δ. Hence

$$\kappa' \geqq \kappa,$$

and this is the relation that we set out to establish. It is unnecessary to consider whether the lower sign can ever hold.

The proof of the lemma is now given by the relation:

$$\int_\alpha^\beta f''(x)^2\,dx > \sum_{i=1}^{n}\int_{\alpha_i}^{\beta_i} f''(x)^2\,dx > \left(\frac{\gamma}{2}\right)^2(\kappa' - \epsilon),$$

from which it follows that

$$\int_\alpha^\beta f'(x)^2\,dx \geqq \tfrac{1}{4}\gamma^2\kappa > \tfrac{1}{8}\frac{\lambda^3}{\tau^2}, \qquad \text{q. e. d.}$$

6. *Proof of the theorem.* — To prove the theorem let us first assume that the curve $y = f'(x)$ cuts the line $y = 1$ in a finite number of points, and that, in the intervals $(a_1, b_1), (a_2, b_2), \cdots, (a_n, b_n)$,

$$f''(x) \leqq 1, \quad a_i \leqq x \leqq b_i,$$

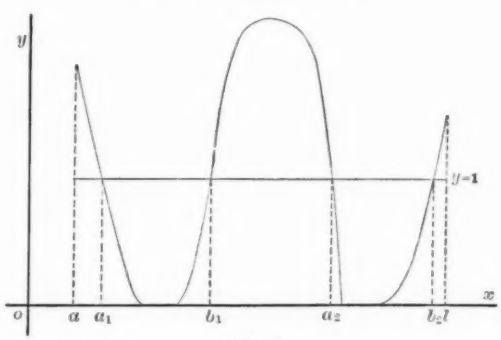

FIG. 6.

while in the remaining intervals, (b_0, a_1), (b_1, a_2), $\cdots$, (b_n, a_{n+1}), where $b_0 = a$, $a_{n+1} = l$,

$$f'(x) > 1, \quad b_i < x < a_{i+1}.$$

(In particular, the curve need not cut this line at all. The result obtained below is still true, and the modification in the proof is but formal.) Observe that we may write down the relation:

$$\sum_{i=0}^{n} [f(a_{i+1}) - f(b_i)] + \sum_{i=1}^{n} [f(b_i) - f(a_i)] = L. \tag{3}$$

The relation on which the proof of the theorem turns is the following:

$$\begin{cases} \displaystyle\int_a^l f'(x)^2\,dx = \sum_{i=0}^{n} \int_{b_i}^{a_{i+1}} f'(x)^2\,dx + \sum_{i=1}^{n} \int_{a_i}^{b_i} f'(x)^2\,dx \\ \qquad\qquad \displaystyle\geqq \sum_{i=0}^{n} [f(a_{i+1}) - f(b_i)] + \tfrac{1}{8} \sum_{i=1}^{n} \frac{[f(b_i) - f(a_i)]^3}{(b_i - a_i)^2}, \end{cases} \tag{4}$$

the approximate evaluation of the first sum being given by the relation $f'(x)^2 > f'(x)$, and that of the second sum by the lemma of paragraph 5. The terms that compose the last sum we divide into two classes, namely:

(A) The terms for which

$$\frac{f(b_i) - f(a_i)}{b_i - a_i} \geqq \frac{1}{\rho} L,$$

where ρ is a positive quantity to be defined later. Let the values of the index for these terms be denoted by $i_1, i_2, \cdots, i_p$;

(B) The remaining terms. For these

$$f(b_j) - f(a_j) < \frac{1}{\rho} L (b_j - a_j);$$

let the values of the index j for these terms be denoted by $j_1, j_2, \cdots, j_m$. Notice that

$$\sum_{k=1}^{m} [f(b_{j_k}) - f(a_{j_k})] < \frac{1}{\rho} L \sum_{k=1}^{m} (b_{j_k} - a_{j_k}) \leqq \frac{1}{\rho} L (b - a). \tag{5}$$

We now write down relation (4), suppressing all terms of class (B) and replacing each summand in class (A) by a smaller quantity, as follows:

$$\frac{[f(b_{i_k}) - f(a_{i_k})]^3}{(b_{i_k} - a_{i_k})^2} = \left\{\frac{f(b_{i_k}) - f(a_{i_k})}{b_{i_k} - a_{i_k}}\right\}^2 [f(b_{i_k}) - f(a_{i_k})] \geqq \frac{L^2}{\rho^2} [f(b_{i_k}) - f(a_{i_k})].$$

The result is the relation

$$\int_a^l f'(x)^2\,dx \geqq \sum_{i=0}^{n} [f(a_{i+1}) - f(b_i)] + \frac{L^2}{8\rho^2} \sum_{k=1}^{p} [f(b_{i_k}) - f(a_{i_k})].$$

Next, let ρ be taken large enough so that $L^2/8\rho^2 < 1$. Then

$$\int_a^l f''(x)^2\,dx > \frac{L^2}{8\rho^2}\left\{\sum_{i=0}^{n}[f'(a_{i+1})-f'(b_i)]+\sum_{k=1}^{\mu}[f'(b_{i_k})-f'(a_{i_k})]\right\}.$$

The quantity in the brace we now transform by adding and subtracting the terms of class (B). The addition of these terms gives three sums whose total value is precisely L. On the other hand, the inequality will only be strengthened if, instead of

$$\sum_{k=1}^{m}[f'(b_{j_k})-f'(a_{j_k})]$$

we subtract the larger quantity given by (5), namely, $\rho^{-1}L(b-a)$. We thus obtain the relation:

$$\int_a^l f''(x)^2\,dx > \frac{L^2}{8\rho^2}\left\{L-\frac{b-a}{\rho}L\right\}.$$

The expression on the right hand side is seen to have its maximum value when $\rho = \frac{3}{2}(b-a)$, and hence it follows that

$$\int_a^l f''(x)^2\,dx > \frac{L^3}{54(b-a)^2},$$

provided that the above is an admissible value for ρ; i. e., if it will make $L^2/8\rho^2 < 1$. This will be the case if

$$L < 3\sqrt{2}\,(b-a),$$

and this relation was made part of the hypothesis of the theorem for this very purpose. Since we are concerned only with small values of L, this restriction is not embarrassing.

We may note in passing that we could obtain a corresponding formula for large values of L by choosing ρ so that $1 - L^2/8\rho^2$ will be an arbitrarily small positive quantity, and then replacing the factor L/ρ in the brace by the larger value $2\sqrt{2}$. The resulting formula (which, indeed, is true for all values of L) is:

$$\int_a^l f''(x)^2\,dx > \{L - 2\sqrt{2}\,(b-a)\}.$$

We have hitherto assumed that the curve $y = f(x)$ meets the line $y = 1$ only in a finite number of points. If this is not the case, then it is possible to divide the interval (a, l) into a finite number of subintervals and to separate these into two classes:

(a) Intervals within which the curve and the line do not meet. They shall meet at each extremity of such an interval, the points $x = a$ and $x = l$ excepted.

(b) Intervals within which the maximum value of $f'(x)$ is less than $1 + \epsilon$, where ϵ is an arbitrary positive quantity.

Let a function $f_1(x)$ be defined as follows:

$f'_1(x) = f'(x)$ in the intervals (a);

$f'_1(x) = f'(x)$ at those points of the intervals (b) at which $f'(x) \leqq 1$;

$f'_1(x) = 1$ at the remaining points.

Then, if

$$f_1(l) - f_1(a) = L_1, \quad L - \epsilon(l - a) < L_1 < L.$$

To the function $f'_1(x)$ all the reasoning is applicable that has been used above in the case of the function $f'(x)$, if we introduce the following modification: The curve $y = f_1(x)$ may be cut by the line $y = 1$ in an infinite number of points. The points of this set in parts of whose neighborhoods $f_1(x)$ is positive are, however, finite in number. These points play the rôle of the points of intersection in the former case; and if we denote them by $a_1, b_1, \cdots a_n, b_n$, the former proof will apply. Thus

$$\int_a^l f'_1(x)^2\, dx \ \geqq \ \frac{L_1^3}{54(b-a)^2}.$$

But

$$\int_a^l f'(x)^2\, dx > \int_a^l f'_1(x)^2\, dx,$$

and we now readily infer that

$$\int_a^l f'(x)^2\, dx \geqq \frac{L^3}{54(b-a)^2}.$$

This completes the proof of the theorem.

§ 3. The Proof of Weierstrass's Theorem.

7. Let the points A, B be joined by an arbitrary curve lying in the field T, but not lying wholly within $\mathfrak{T}$:

$$C: \qquad x = \phi(t), \quad y = \psi(t); \qquad t_0 \leqq t \leqq t_1,$$

where ϕ, ψ denote, to begin with, merely single-valued, continuous functions of

t, and let a polygon P be inscribed in C with its vertices at the points $\tau_0 = t_0$, $\tau_1, \cdots, \tau_n = t_1$, where $\tau_i < \tau_{i+1}$. Its sides may intersect one another. Form the sum

$$S = \sum_{i=0}^{n-1} F(x_i, y_i, \Delta x_i, \Delta y_i).$$

Let the curve C be now restricted to belonging to the class (K) of curves for which S approaches one and the same limit when $n = \infty$, no matter how the vertices of the polygon are chosen, provided merely that its longest side converges towards 0. Denote this limit by $\mathfrak{J}$. Then WEIERSTRASS'S theorem is established if we show that

$$\mathfrak{J} \geqq J + \frac{L^3}{54(u_1 + L - u_0)^2}, \tag{1}$$

and this we will now do.

Let ϵ and δ be two arbitrarily chosen positive quantities, and let the polygon P be so chosen that

$$|\mathfrak{J} - S| < \epsilon \tag{2}$$

and that the length of the longest side of P is less than δ. Form the (improper) integral I for the curve $C = P$. This will be a proper integral provided we assign to the arguments x', y' definite (but arbitrary) values at the vertices of P. We have then:

$$\begin{aligned} I = \int_{t_0}^{t_1} F(x, y, x', y')\,dt &= \sum_{i=0}^{n-1} \int_{\tau_i}^{\tau_{i+1}} F(x, y, x', y')\,dt \\ &= \sum_{i=0}^{n-1} F\left(\bar{x}_i, \bar{y}_i, \frac{\Delta x_i}{\Delta t_i}, \frac{\Delta y_i}{\Delta t_i}\right)\Delta t_i = \sum_{i=0}^{n-1} F(\bar{x}_i, \bar{y}_i, \Delta x_i, \Delta y_i), \end{aligned} \tag{3}$$

where

$$\bar{x}_i = \phi(t_i + \vartheta_i \Delta t_i) = \phi(t_i) + \zeta_i, \qquad 0 < \vartheta_i < 1,$$

and

$$\bar{y}_i = \psi(t_i + \vartheta_i' \Delta t_i) = \psi(t_i) + \zeta_i', \qquad 0 < \vartheta_i' < 1.$$

By proper choice of δ, it is possible to make

$$|\zeta_i| < h, \quad |\zeta_i'| < h \qquad (i = 0, 1, \cdots, n-1), \tag{4}$$

where h denotes an arbitrary positive constant.

The function $F(x, y, \cos\phi, \sin\phi)$ is uniformly continuous and hence, by proper choice of h, we can insure that

$$|F(x_i + \zeta_i, y_i + \zeta_i', \cos a_i, \sin a_i) - F(x_i, y_i, \cos a_i, \sin a_i)| < \epsilon, \tag{5}$$

where $\Delta x_i / p_i = \cos a_i$, $\Delta y_i / p_i = \sin a_i$ and $p_i = \sqrt{\Delta x_i^2 + \Delta y_i^2}$. Multiplying (5)

through by p_i and adding from $i=0$ to $i=n-1$, we get, remembering that $F(x_i, y_i, \cos a_i, \sin a_i)p_i = F(x_i, y_i, \Delta x_i, \Delta y_i)$ etc.,

$$|I - S| < \epsilon l_a. \tag{6}$$

Since $F(x, y, \sin\phi, \cos\phi)$ is always positive and is continuous, it has a positive lower limit when x, y, ϕ vary arbitrarily, and hence it is only for curves that have a length that S can converge toward a limit, as is at once evident if the arc S of P is used as the parameter t in (3). Denoting the length of C by l, we replace (6) by the relation:

$$|I - S| < \epsilon l. \tag{7}$$

From (2) and (7) it follows that

$$|\mathfrak{I} - I| < \epsilon(l+1).$$

But, from the theorem of § 1,

$$I \geqq J + \frac{L^3}{54(u_1 + L - u_0)}$$

and hence

$$\mathfrak{I} > J + \frac{L^3}{54(u_1 + L - u_0)^2} - \epsilon(l+1).$$

Since ϵ is arbitrary, the truth of the relation (1) is thus established.*

* Note of July 6. HILBERT has given a definition of the integrals

$$\int F(x, y, y')\,dx \qquad \iint F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right)dx\,dy,$$

which, like the Weierstrassian definition of the former of these (written in homogeneous form), is a generalization of the ordinary definition for the case that the function y or z does not have a continuous first derivative or continuous first partial derivatives; cf. E. R. HEDRICK, *Ueber den analytischen Character der Lösungen von Differentialgleichungen*, Dissertation, Göttingen, 1901, p. 69. For the integral considered in the present section, HILBERT's definition is in substance the following. Let the continuous curve C be divided as before into n segments by the points $\tau_0 = t_0, \tau_1, \cdots, \tau_n = t_1$, where $\tau_i < \tau_{i+1}$. Through these n points pass an arbitrary curve, which HILBERT assumes to be analytic; it is enough to assume that it is made up of a finite number of pieces, each of which is a continuous curve having a continuous tangent throughout its whole extent, its extremities included. From the integral $I = \int_{t_0}^{t_1} F(x, y, x', y')\,dt$ for this curve. Then when all such curves through the $n+1$ points are considered, the various values of the integral will have a *lower limit*. In the case of the geodesic integral this lower limit is obviously the length of the inscribed polygon above considered. Next, consider the totality of values consisting of these lower limits; to each one of the infinity of modes of division of C into n parts corresponds one such value, and $n = 1, 2, \cdots$. The *upper limit* of these values HILBERT defines as the value of the integral in question.

Thus we have, in form at least, a new generalization of the integral I for curves C for which this upper limit is finite; and the totality of such curves form a class (K'). *This class coincides, however, with the class* (K) *and the value of the generalized integral for a curve of the*

new class coincides with the value of the integral generalized according to Weierstrass's definition for the same curve. The proof of this theorem depends on the fact that the integral I can be made a minimum by a curve c connecting any two points of T, provided that the distance between these points does not exceed a suitably chosen positive constant h, and that these curves have continuous curvature which lies (numerically) uniformly below a suitably chosen positive constant G. Thus, when n is large and the maximum distance between two successive points τ_i, τ_{i+1} is small, the lower limit of the integral will be obtained by employing the curve made up of the n curves connecting the $n+1$ points, each curve being so chosen as to make I a minimum over against all other curves having the same extremities. These curves differ but slightly from right lines in all of their properties that are essential for the above proof.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS., *April*, 1901.

CONCERNING HARNACK'S THEORY OF IMPROPER DEFINITE INTEGRALS*

BY

ELIAKIM HASTINGS MOORE

INTRODUCTION.

In this paper I consider the improper simple definite integrals of HARNACK (1883, 1884). In the introduction I wish to characterize somewhat clearly the theories of the improper simple and multiple integrals recently given by JORDAN (1894) and STOLZ (1898, 1899), and in this introductory paragraph I summarize the contents of the whole introduction. These theories for the simple integrals have intimate relations with the HARNACK theory. The definition adopted for the multiple integrals is more exacting than that for the simple integrals. The multiple integrals converge or exist (as limits) only absolutely. For the simple integrals we have then two theories, on the one hand, of the integrals with the milder definition, and, on the other hand, of the integrals with the stronger definition and so with a larger body of properties. The first class of integrals includes the second class of integrals. The HARNACK theory relates to the first and general class of integrals; this theory has not received systematic development; however, for the theory of the absolutely convergent HARNACK integrals this is not true, and these integrals constitute the second and special class of integrals. I discuss both classes of simple integrals simultaneously and by uniform process; this is made possible by suitable determinations of the definitions; the absolute convergence of the integrals of the second class appears only at the conclusion, and hence it is desirable to introduce terms of discrimination connoting the two definitions, the milder and the stronger; the terms chosen, "narrow," "broad," connote the geometric form of the definitions, and likewise the fact that the class of narrow integrals has a less extensive body of properties than the (included) class of broad integrals. There has been a tendency to do away with the non-absolutely convergent HARNACK integrals; I hope to show that this tendency rests upon misconceptions.—The theory of DE LA VALLÉE POUSSIN (initiated in 1892) is in form distinct from the HARNACK theory and

* Presented to the Society (Chicago) December 28, 1900. Received for publication June 12, 1901.

relates primarily to absolutely convergent integrals. It is not involved in the present paper.

HARNACK, in papers* published in 1883 and 1884, in volumes 21 and 24 of the Mathematische Annalen, first gave a definition for the notion of an improper definite integral from a to b of a real function $F(x)$ having these two properties:

(1) The function $F(x)$ assumes values indefinitely great in the neighborhood of points ζ constituting a point-set† Z lying on the interval ab and of content zero; and

(2) The function $F(x)$ is properly integrable in the neighborhood of every other point x of ab, or what is equivalent, it is properly integrable from a' to b' where $a'b'$ is any interval of ab, which contains no point ζ.

HARNACK's definition was formally a definition not so much of the definite integral

$$\int_a^b F(x)\,dx \tag{1}$$

itself as of the definite integral function

$$J(X) = \int_a^X F(x)\,dx \qquad (X \text{ of } ab).$$

Indeed his definition, in common with many limit-definitions of that period and of earlier periods, was not expressed in explicit form. From the context one may infer that HARNACK was so desirous of passing to the less immediate applications of the notion that he was unwilling to attend to the systematic exposition of the fundamental elements of the theory. Thus he was led into error,—with respect to the general theory, notably in the theorem‡:

If the set Z of singularities of the function $F(x)$ is reducible (and so of content zero), and a continuous function $\phi(x)$ of x on ab exists, for which

$$\phi(x'') - \phi(x') = \int_{x'}^{x''} F(x)\,dx$$

for every interval $x'x''$ of ab, containing no singularity ζ, then the integral

---

**Anwendung der Fourier'schen Reihe auf die Theorie der Functionen einer complexen Veränderlichen;* vol. 21 (1883), pp. 305–326; cf. pp. 324–326. *Die allgemeinen Sätze über den Zusammenhang der Functionen einer reellen Variabelen mit ihren Ableitungen, II. Theil;* vol. 24 (1884), pp. 217–252; cf. p. 220 fg.

† It is convenient to use *set* as the equivalent of *Menge* and *ensemble*.—For CANTOR's *Punktmenge* OSGOOD has used the English term *Cantor's set*.

‡ Loc. cit., vol. 24, p. 222, theorem 3. This theorem depends on theorem 2, which is likewise erroneous (cf. § 5 6°).

$$\int_a^x F(x)\,dx \qquad (x \text{ of } ab)$$

exists * or converges and has the value

$$\phi(x) - \phi(a).$$

This theorem was designed to show that the improper definite integrals defined by HÖLDER (Mathematische Annalen, vol. 24 (1884), p. 190 fg.) are special cases of the HARNACK integrals, the set Z being reducible. And this does follow, by the elucidations of HÖLDER as to the uniqueness of his definite integrals, whenever the corresponding HARNACK integrals exist. Thus the HARNACK integrals with reducible sets of singularities are HÖLDER integrals, and the HÖLDER integrals for which the corresponding HARNACK integrals (with reducible sets of singularities) exist are these HARNACK integrals. By examples in § 5 (4°, 9°) I show that each system of integrals is more extensive than their common subsystem.

I accept the sharp formulation of the limit-definition of the integral (1) given by STOLZ† (1898) as doubtless a correct expression of the content of HARNACK's meaning.

As SCHOENFLIES‡ has remarked, the generalized simple definite integrals of JORDAN (*Cours d'analyse*, ed. 2, vol. 2 (1894), pp. 46, 50 fg.) are§ HARNACK's improper integrals, although there is no reference to HARNACK and the form of exhibition is different. JORDAN (loc. cit., p. 46) selects two properties of the

* I think of the integral as a limit, that is, as a certain number obtained by a certain limiting process, and prefer to say in general that the limit exists rather than that it converges; the limit-and expression converges to the number in question as its limit.

† The full reference is given below.

‡ Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 8: *Bericht* (1900), p. 176.—SCHOENFLIES (loc. cit., p. 186) gives suggestively but not explicitly a definition of the HARNACK integrals in terms involving the set of non-overlapping intervals of ab which enclose no point ζ and every point not-ζ enclosed by ab. One may form three explicit definitions in these terms. The third definition agrees with the definition given by STOLZ, and the second is equivalent to it, while the first, which alone fully expresses the implications to me of the language used by SCHOENFLIES, is a definition considerably milder.

§ At least if the content of the set Γ of points c (loc. cit., p. 50) is zero. SCHOENFLIES suggests that a condition to this effect may have been unintentionally omitted. But JORDAN defines (loc. cit., p. 76) double integrals without the insertion of the analogous condition; for (contrary to a statement of SCHOENFLIES) he speaks merely of the *interior* content of the region of integration. However, in the case of the simple integrals, if this condition is omitted, the theorem cited in the text will fail, whenever the condition for $\int_a^b F(x)\,dx$ is not satisfied, even for all cases in which $G(x)$ is a non-zero constant. It may be noted here that JORDAN's set Γ is the set of all points c in whose neighborhood the function $F(x)$ is not capable of proper integration. Thus, the set Γ includes HARNACK's set Z. But we have the theorem that the set Γ is HARNACK's set Z, in case Γ has content zero and the improper definite integral exists. Cf. theorem VIIA of § 3.

proper integrals which serve to suggest the explicit definition of the generalized integrals. But the theorems of the sequel and the reference theorem implied (p. 226, ll. 9, 10) in the proof of the second mean value theorem for the generalized integrals, certainly need fuller proofs: indeed, the theorem* that $\int_a^b \big(F(x)+G(x)\big)\,dx$ exists and

$$\int_a^b \big(F(x)+G(x)\big)\,dx = \int_a^b F(x)\,dx + \int_a^b G(x)\,dx$$

in case the integrals on the right exist, is not in general true. In § 5 8° I give a simple example in which the integral in question does not exist.

Quite recently HARNACK's theory (without reference to JORDAN's theory) has been considered systematically and critically by STOLZ, first in 1898 in a paper† entitled "*Zur Erklärung der absolut convergenten uneigentlichen Integrale*," and then in 1899 (with but slight modifications) in appendix III to volume 3 of his *Grundzüge der Differential- und Integralrechnung*, and further immediately thereafter, in 1899, in a paper entitled "*Über die absolute Convergenz der uneigentlichen bestimmten Integrale*."†

STOLZ gives‡ an explicit definition of the HARNACK integral (1) as a certain limit, and then, considering the theory of these integrals, he affirms‡ that it fails to justify for the limit in question the notation and designation *definite integral* —that one has not even the property that if $\int_a^b F(x)\,dx$ exists, so do all the integrals $\int_a^c F(x)\,dx$ $(a<c<b)$. Contenting himself at this point with an affirmation, STOLZ turns to the theory of the absolutely convergent integrals (1), that is, those for which also the corresponding integral $\int_a^b |F(x)|\,dx$ exists, and finds§ that these integrals deserve the name, since they possess certain four fundamental properties.

The definition given by STOLZ has application with respect to any point-set Ξ lying on the interval ab and of content zero, which contains the set Z used by HARNACK, and in so far his definition is formally more comprehensive than HARNACK's definition. This extension was necessary for the formulation of the third and fourth of the four fundamental propositions which are as follows:

Under the hypothesis that $F(x)$ is absolutely integrable from a to b with respect to a point-set Ξ of content zero:

I. $F(x)$ is likewise integrable from a to c and from c to b, where c is any point of ab, and

---

* Theorem 3, § 58, p. 56.—This theorem would follow directly from HARNACK's erroneous theorem 2 cited above.

† Wiener Berichte, vol. 107_{IIa}, pp. 207–224; vol. 108_{IIa}, pp. 1234–1238.

‡ Wiener Berichte, vol. 107, pp. 207, 211. *Grundzüge*, vol. 3, p. 277.

§ Wiener Berichte, vol. 107, p. 211, § 2, and pp. 216–221, § 5. *Grundzüge*, vol. 3, pp. 277, 279–284.

$$\int_a^b F(x)\,dx = \int_a^c F(x)\,dx + \int_c^b F(x)\,dx.$$

II. For every positive ϵ there is a positive δ_ϵ such that for any set of a finite number (n) of non-overlapping intervals $a_k b_k$ $(a_k < b_k;\ k = 1, \cdots, n)$ of the interval ab of total length less than δ_ϵ the sum of the n integrals $\int_{a_k}^{b_k} F(x)\,dx$ is in absolute value less than ϵ.

III. $F(x)$ is absolutely integrable from a to b with respect to the point-set $\Xi + \mathrm{H}$ obtained by extending Ξ by any set H of content zero, and the integrals with respect to Ξ and to $\Xi + \mathrm{H}$ are equal.

IV. If $G(x)$ is absolutely integrable from a to b with respect to the point-set H of content zero, then $F(x) + G(x)$ is likewise integrable with respect to the aggregate set $\Xi + \mathrm{H}$, and

$$\int_a^b \big(F(x) + G(x)\big)\,dx = \int_a^b F(x)\,dx + \int_a^b G(x)\,dx.$$

As to these theorems it is to be noted that the third is an immediate consequence of the first two, and that the fourth is an immediate consequence of the third.

As to the corresponding theorems for the HARNACK integrals in general I shall show (§ 3 V, VIII) that the first is in fact true, notwithstanding the statement* of STOLZ to the contrary; the second is the erroneous theorem 2 of HARNACK and the fourth is the erroneous theorem 3 of JORDAN. Indeed, the validity of the second theorem for a HARNACK integral implies its absolute convergence (cf. § 2, def. 1, note 1, footnote, and § 4 III). The second theorem is in effect a generalization of the uniform continuity of the definite integral function $J(x)$; for the non-absolutely convergent HARNACK integrals the uniform continuity of $J(x)$ holds (§ 3 IX, § 5 6°).

The limit-definition of the integral (1) relates to the various interval-sets I of a finite number of intervals enclosing all points ζ or ξ of the set Z or Ξ; in notation $I = I(\mathrm{Z})$ or $I(\Xi)$. According to the implication of HARNACK and of STOLZ and the explicit statement of JORDAN (loc. cit., p. 51, top) every interval of I encloses (at least) one point ζ or ξ.

JORDAN'S definitions† for the improper simple and multiple integrals relate to closed measurable regions D lying within the bounded or limited region E

* This statement of STOLZ is reaffirmed by PRINGSHEIM (Münchener Berichte, 1900, p. 220) in his recent paper on the second mean value of the integral calculus, and it is the basis of PRINGSHEIM'S rejection of HARNACK'S extension (Annalen, vol. 21, p. 326) of that theorem to the general improper integrals; the theorem as extended is, however, correct, and the following development of the general theory makes HARNACK'S elegant proof effective.

† loc. cit., pp. 51, 76, and 87. The theory of the improper upper and lower multiple integrals is first developed.

of integration and containing no singularity c. For the simple integrals the complementary region $E - D$ is the interval-set $I(\Xi) = I(\Gamma)$ each of whose intervals encloses a point c. For the multiple integrals the corresponding condition is not imposed, and accordingly the condition for the existence of the improper integral is stronger, and JORDAN proves* that the multiple integrals as defined exist only absolutely. For the simple integrals the corresponding theorem does not hold, and JORDAN (p. 87) remarks that the loss of the theorem is due to the fact that the definition for the simple integrals is less exacting, although he expresses himself somewhat obscurely, in terms of one dimension not capable of immediate extension to more dimensions.

STOLZ† develops the JORDAN definition of improper double integrals for the case‡ of a region E with ordinary boundaries on which all the singularities c lie. The one-dimensional analogue of this case is that of a finite interval E with singular extremities, and for the simple integrals D is a sub-interval of E. In his last paper§ STOLZ, following WIRTINGER, notices that if (as in two dimensions) the D is allowed to be an interval-set the improper integrals exist only absolutely. (This, we have seen, was the implication of the remark of JORDAN.) And, further, he likewise notices (as a generalization for the special case of the theorem of JORDAN) that the double integrals would still exist only absolutely even if they were given a less exacting definition, the region D being required to be (as in the case of the simple integrals) connected‖ or of one piece.

To revert to the general case of multiple integrals, it is now apparent that for a definition formally less exacting than JORDAN's we may impose on the region D which converges to E the conditions: (1) the region $E - D$ consists of one or more regions, each being of one piece and each enclosing a singular point c; (2) the region D consists of one or more regions, each being of one piece and no two being capable of union as parts of an including connected region D; (3) the region D consists of a finite number of connected regions. And these conditions on the region D may be imposed or not imposed independently of one another.

In this paper I confine attention to the simple integrals over a finite interval $E = ab$; the conditions (1) and (2) then become identical; we impose the condition (3), and thus have the theory of the narrow or of the broad simple inte-

* loc. cit., pp. 80, 87.

† *Grundzüge*, vol. 3 (1899), p. 122 fg.

‡ This case is hardly equivalent to the general case, although STOLZ considers that it is (cf. loc. cit., p. 122). The proof (p. 141 fg.) of the theorem of the absolute existence of the improper double integrals I find inconclusive.

§ Wiener Berichte, vol. 108 (1899), pp. 1234–1238.

‖ SCHOENFLIES (loc. cit., pp. 203–205) in his report on this paper of STOLZ seems to confuse definitional conditions for the existence of an integral and properties resulting from its existence.

grals, according as the interval-set $E - D$ is not or is required to have on its every interval a singularity c. By developing the two theories simultaneously we obtain a new insight into the varying properties of the general and the absolutely convergent HARNACK integrals.

In my judgment the general integrals—even the non-absolutely convergent integrals—deserve to be classed with the proper definite integrals under the designation definite integrals. To be sure they lack in general many fundamental properties of the absolutely convergent integrals. We may however look at these matters in such a way that the properties of the different classes are more nearly equivalent. Let us think not of the proper or improper definite integral but, more generally, of the definite integral with respect to a point-set Ξ of content zero and thus write for instance $\int_{a\,\Xi}^{b} f(x)\,dx$, and speak of the Ξ-integrals. Then for every Ξ there is a theory of the narrow and of the broad Ξ-integrals, and for the various sets Ξ these theories, in so far as they relate to a single set Ξ or to a single set Ξ and its subsets Ξ_0 including the set Z of singularities of the integral function $F(x)$, are to a large extent the same. (I remark in passing that the desirable theorem (§ 2 V): if $F(x)$ is Ξ-integrable from a to b and the set Z is non-existent, then $F(x)$ is properly integrable from a to b and $\int_a^b F(x)\,dx = \int_{a\,\Xi}^{b} F(x)\,dx$, holds only if the content of the set Ξ is zero. And in this fact I see one of the strongest reasons for considering only such sets Ξ.)

In the sequel the theories of the narrow and the broad HARNACK integrals are developed from this point of view and otherwise essentially in the spirit of the original HARNACK papers. The necessary fixed hypotheses, etc., being introduced in § 1, I give in § 2 for the Ξ-integrals the definitions and a number of fundamental theorems, and then in § 3 I develop the properties of a function $F(x)$ Ξ-integrable from a to b. I bound the inquiry of § 3 in effect by the form of the four fundamental theorems of STOLZ. Then in § 4 the questions of absolute convergence are introduced. And in § 5 (added July 1, 1901) after exhibiting an important condition necessary and sufficient for the existence of the general or narrow HARNACK integral, I construct for the general closed point-set Ξ on ab of content zero a Ξ-integral from a to b for which the set Ξ is the set Z of singularities, and which is near every point $x = \xi$ essentially narrow or non-absolutely convergent; and in connection with the simplest case, $\Xi = (b)$, I exhibit the examples needed to show the error of various statements already referred to.

§ 1.

FUNDAMENTAL HYPOTHESES, DEFINITIONS, AND NOTATIONS.

1°. We consider the finite interval ab of values of the real variable x. In the proofs of the propositions we shall consider the case $a < b$, to which the case $a > b$ is immediately reducible.

2°. We consider further a point-set Ξ of points ξ which is of content zero. As explained in § 2 the set Ξ will later be supposed to be closed; this supposition involves no restriction of generality. The aggregate or sum $\Xi_1 + \Xi_2$ of two sets Ξ_1, Ξ_2, each closed and of content zero is again a set of the same kind.

3°. On the interval ab the real function $F(x)$ is supposed to be single valued wherever defined, and it is supposed to be defined certainly at all points x not of Ξ. Moreover it is supposed that $F(x)$ is capable of proper definite integration from a' to b' where the interval $a'b'$ is any interval lying on the interval ab and containing no point ξ and no limit-point ξ' of Ξ.

4°. With respect to $F(x)$ it is convenient to separate points $x = x_0$ into two classes. The point $x = x_0$ is *regular* if near x_0 (i. e., on some interval $x_0 - \delta \cdots x_0 + \delta$) $F(x)$ is everywhere defined, single valued, and limited; and otherwise it is *singular*. A singular point $x = \zeta$ may be both progressively and regressively singular or merely progressively ($\zeta = \zeta_+$) or regressively ($\zeta = \zeta_-$) singular. We speak of the singularities ζ and of the singular point-set Z of all points ζ. The set Z is closed. Obviously every singular point ζ is a point ξ or a limit-point ξ' of Ξ; if the set Ξ is closed, the set Z is a subset of it.

5°. A finite number of intervals such that no two have a common point is called an interval-set. Denoting by I an interval-set we denote its length, the sum of the lengths of its intervals, by D_I.

6°. Two intervals $i_1 i_2$ having a common inner point determine a definite interval $i_{12} = i_{21}$ common to i_1 and i_2. Two interval-sets $I_1 I_2$ having a common inner point determine a definite interval-set $I_{12} = I_{21}$ which may be called the set of intervals common to $I_1 I_2$, that is, every i_{12} of I_{12} is the interval common to certain two intervals i_1 of I_1, i_2 of I_2. The set common to a set I and an interval $x_1 x_2$ is denoted by $I_{x_1}^{x_2}$.

For brevity, especially in partitioning interval-sets, it is sometimes convenient to use these notations, even when the interval-sets in question do not exist; for example, an interval-set I is, by an interval $x_1 x_2$ separated into two interval-sets $I_{x_1}^{x_2}$, I' (of which one may not exist), and we write $I = I_{x_1}^{x_2} + I'$.

7°. An interval-set I is said to contain any point-set S, all of whose points are points of I; and in this case S is said to be of or to lie on I. Further an interval-set is said to enclose a point-set S if every point s and likewise every limit-point s' of S lies within some interval of I; and in this case S is said to be an inner set of or to lie within I. The interval-set I is said to contain or to enclose the point-set S narrowly, and S to lie respectively on or within I narrowly, if furthermore every interval of I contains or encloses at least one point s of S. It is convenient to denote by $I(S)$ an interval-set which encloses S narrowly, and by $I\{S\}$ one which encloses S not necessarily narrowly, or, say, one which encloses S broadly. Thus the broad is the generic enclosure, and the narrow is a specific enclosure.

8°. The set Ξ being by hypothesis of content zero, for any positive number ϵ there an interval-set I_ϵ enclosing Ξ narrowly and of length $D_{I_\epsilon} < \epsilon$.

9°. On the fixed interval ab in connection with any interval-set I (not necessarily lying on the interval ab) we introduce a function $F_I(x)$ by the following stipulation: according as the point x of ab lies or does not lie on I, $F_I(x)$ has the value 0 or $F(x)$, with the understanding that, if I is the symbol of a non-existent interval-set, $F_I(x) = F(x)$ for every x of ab. It is to be noticed that if I encloses Ξ, then $F_I(x)$ is on the interval ab everywhere defined, single valued, limited, and capable of proper definite integration from a to b.

§2.

THE DEFINITE INTEGRALS NARROW AND BROAD

$$\int_{a\,(\Xi)}^{b} F(x)\,dx, \qquad \int_{a\,\{\Xi\}}^{b} F(x)\,dx$$

OF THE FUNCTION $F(x)$ FROM a TO b WITH RESPECT TO THE SET Ξ: DEFINITIONS, CONDITIONS NECESSARY AND SUFFICIENT FOR EXISTENCE, AND FUNDAMENTAL THEOREMS.

We consider simultaneously the two cases:

(1) the narrow Ξ-integral, $\int_{a\,(\Xi)}^{b} F(x)\,dx$;

(2) the broad Ξ-integral, $\int_{a\,\{\Xi\}}^{b} F(x)\,dx$.

When these integrals are considered simultaneously and disjunctively (as in §§ 2, 3) we speak simply of

(3) the Ξ-integral: $\int_{a\,\Xi}^{b} F(x)\,dx$.

In case I denotes an interval-set enclosing Ξ narrowly or broadly the function $F_I(x)$ is capable of proper definite integration from a to b. We consider the proper definite integrals:

(4) $$\int_a^b F_I(x)\,dx,$$

for the various interval-sets I enclosing Ξ $^{\text{narrowly}}_{\text{broadly}}$, and denote by the respective limit-notations:

(5) $$\lim_{D_I=0}^{I\ I\,(\Xi)} \int_a^b F_I(x)\,dx; \qquad \lim_{D_I=0}^{I\ I\,\{\Xi\}} \int_a^b F_I(x)\,dx,$$

certain finite constant limits, if existent. Then we give those limits the specific Ξ-integral notations:

$$\int_{a\,(\Xi)}^{b} F(x)dx; \quad \int_{a\,\{\Xi\}}^{b} F(x)\,dx. \tag{6}$$

Thus we have the explicit

Definition * 1. *The (existent)* $\overset{narrow}{broad}$ Ξ*-integral:*

$$\int_{a\,(\Xi)}^{b} F(x)\,dx = \lim_{D_I=0}^{I\ I\ (\Xi)} \sum \int_{a}^{b} F_I(x)\,dx;$$

$$\int_{a\,\{\Xi\}}^{b} F(x)\,dx = \lim_{D_I=0}^{I\ I\ \{\Xi\}} \sum \int_{a}^{b} F_I(x)\,dx; \tag{7}$$

is a certain finite constant such that for every positive number ϵ *there exists a positive number*† δ_ϵ *such that*

$$\left| \sum \int_{a}^{b} F_I(x)\,dx - \int_{a\,\Xi}^{} F(x)\,dx \right| < \epsilon \tag{8}$$

for every interval-set‡ I *enclosing* Ξ $\overset{narrowly}{broadly}$ *and of length* $D_I < \delta_\epsilon$.

Note 1. The broad Ξ-integrals constitute a special type of the narrow Ξ-integrals §. In this connection the adjectives *narrow* and *broad* may connote the fact that the body of properties of the narrow integrals is less extensive than and included in the body of properties of the broad integrals. The essentially narrow integrals are the not-broad narrow integrals.

---

* This definition (of the narrow Ξ-integral) seems to express (for $\Xi = Z$) exactly HARNACK'S meaning in form as well as in content.

HARNACK defines the improper definite integral over a certain interval as a certain limit of the proper definite integral over the same interval of a modified function. There are certain advantages (cf. the remark of § 3 VI) in this type of definition. A different definition of this type is that of DE LA VALLÉE POUSSIN for the absolutely convergent generalized definite integrals. (With respect to this matter SCHOENFLIES (loc. cit., pp. 186, 187) has erred in setting in contrast the two limiting processes in question.)

† To the various δ's related in this paper to the arbitrary positive ε I give the notations $\delta_\epsilon, \delta'_\epsilon, \delta^1_\epsilon, \delta^2_\epsilon$, etc., where the superscripts are discriminating affixes and not exponents.

‡ This fundamental definitional property of a function $F(x)$ Ξ-integrable from a to b is (with modification of the δ_ϵ) considerably extended in theorem XIV of § 3.

§ That the narrow Ξ-integral (1) be a broad Ξ-integral (2) it is obviously necessary and sufficient that for every ε a δ^3_ϵ exist such that, in the notation (11),

$$\left| \int_J^* F(x)\,dx \right| < \varepsilon$$

for every interval-set J lying on ab and containing no point ξ and of length $D_J < \delta^3_\epsilon$. Indeed this condition (cf. JORDAN, loc. cit., § 74, p. 77) is necessary and sufficient for the existence of the broad integral (2). The reader will compare this remark with § 3 XIII′.

Note 2. From the definition it is apparent that the Ξ-integral and the $(\Xi + \Xi')$-integral, where Ξ' is the set of all limit-points ξ' of Ξ, coexist and are equal; that is, if either exists, so does the other with the same value.—The set $\Xi + \Xi'$ is closed and of content zero.—*We suppose hereafter that the set* Ξ *is closed.* This supposition involves no essential loss of generality and it facilitates the phrasing of many proofs.

If this Ξ-integral exists we say that $F(x)$ is Ξ-integrable from a to b.

Definition 2. *The function* $F(x)$ *is progressively* $\overset{\text{narrowly}}{\text{broadly}}$ Ξ*-integrable at* $x = x_0$ *if there is an interval* $x_0 \cdots x_0 + \epsilon\,(\epsilon > 0)$ *such that the* $\overset{\text{narrow}}{\text{broad}}$ Ξ*-integral from* x_0 *to* $x_0 + \epsilon$ *exists.—Regressive* Ξ*-integrability at a point is similarly defined.*

Definition 3. *The function* $F(x)$ *is* $\overset{\text{narrowly}}{\text{broadly}}$ Ξ*-integrable on* ab $(a < b)$ *if at* a *it is progressively, at* b *it is regressively, and at every other point of* ab *it is both progressively and regressively* $\overset{\text{narrowly}}{\text{broadly}}$ Ξ*-integrable.*

Remark. It will appear (§ 3, theorem V, corollary) that if $F(x)$ is Ξ-integrable on ab it is Ξ-integrable from a to b.

One proves by the usual limit-considerations the following two theorems:

I. *The* Ξ*-integral if existent is uniquely existent.*

II. *For the existence of the definite* $\overset{\text{narrow}}{\text{broad}}$ Ξ*-integral*

$$\int_{a\,\Xi}^{b} F(x)dx \tag{9}$$

it is necessary and sufficient that for every ϵ *a* δ_ϵ^1 *exists such that*

$$\left| \int_a^b F_{I_1}(x)dx - \int_a^b F_{I_2}(x)dx \right| < \epsilon \tag{10}$$

for every pair of interval-sets $I_1 I_2$ *each enclosing* Ξ $\overset{\text{narrowly}}{\text{broadly}}$ *and of length less than* δ_ϵ^1.

Remark. If the Ξ-integral (9) exists, the number $\delta_{\epsilon/2}$ is effective as a number δ_ϵ^1 of the necessary condition (10). And if the sufficient condition (10) is fulfilled the number $\delta_{\epsilon/2}^1$ is effective as a number δ_ϵ connected with the conclusion that the Ξ-integral (9) exists.

III. *If the function* $F(x)$ *is* Ξ*-integrable from* a *to* c *and from* c *to* b, *where* c *lies between* a *and* b, *then the function* $F(x)$ *is* Ξ*-integrable from* a *to* b, *and*

$$\int_{a\,\Xi}^{b} F(x)dx = \int_{a\,\Xi}^{c} F(x)dx + \int_{c\,\Xi}^{b} F(x)dx.$$

One proves this theorem by use of the corresponding theorem for proper definite integrals together with theorem I. The converse of this theorem is theorem VIII of § 3.

Notation. For a function $f(x)$ properly integrable on the various intervals $a_l b_l$ $(a_l < b_l;\ l = 1, 2, \cdots, n)$ of an interval-set J we introduce the notation

$$\int_J f(x)\,dx, \tag{11}$$

by the definitional equation:

$$\int_J f(x)\,dx = \sum_{l=1}^{n} \int_{a_l}^{b_l} f(x)\,dx; \tag{12}$$

and, in case J is the symbol of a non-existent interval-set, it is convenient to understand by the notation (11) the constant zero.

Lemma I. If the function $f(x)$ is properly integrable from a to b and J is any interval-set of the interval ab, then

$$\left|\int_J f(x)\,dx\right| < D_J\, C,$$

where C is any positive number greater than $|f(x)|$ for every point x of the interval ab.

This is a fundamental theorem of the theory of proper definite integrals.

Lemma II. If the function $f(x)$ is properly integrable from a to b, then for every ϵ there is a δ'_ϵ such that

$$\left|\int_J f(x)\,dx\right| < \epsilon$$

for every interval-set J lying on ab and of length $D_J < \delta'_\epsilon$.

Lemma II follows from lemma I immediately. Lemma II is a generalization of the theorem of the uniform continuity on the X-interval ab of the definite integral function $\int_a^X f(x)\,dx$.

IV. *If the function $F(x)$ is properly integrable from a to b, then with respect to every set Ξ of content zero the function $F(x)$ is Ξ-integrable from a to b, and*

$$\int_{a\,\Xi}^{b} F(x)\,dx = \int_a^b F(x)\,dx.$$

*Corollary.** If the function $F(x)$ is properly integrable from a to b and a function $G(x)$ is defined and equal to $F(x)$ on the interval ab, except perhaps at points of a set Ξ of content zero, then $G(x)$ is Ξ-integrable from a to b, and

$$\int_{a\,\Xi}^{b} G(x)\,dx = \int_a^b F(x)\,dx.$$

* Compare the remark of STOLZ, *Grundzüge*, vol. 3 (1899), pp. 283, 284.

Taking a definite set Ξ we denote by I any interval-set enclosing Ξ $\begin{matrix}\text{narrowly}\\ \text{broadly}\end{matrix}$; by J the interval-set I_a^b, and by K the interval-set making up with J the interval ab. Either J or K may be non-existent, but one at least is existent. In these notations we have evidently

$$\int_a^b F(x)\,dx = \int_J F(x)\,dx + \int_K F(x)\,dx,$$

$$\int_a^b F_I(x)\,dx = \int_K F(x)\,dx,$$

and so

$$\int_a^b F_I(x)\,dx - \int_a^b F(x)\,dx = -\int_J F(x)\,dx.$$

From this equality, since $D_J \leqq D_I$, the truth of theorem IV appears by the use of lemma II, the definition of the Ξ-integral, and theorem I.

V. *If the function $F(x)$ exists on the interval ab as a single valued limited function of x, and if it is with respect to a certain set Ξ of content zero Ξ-integrable from a to b, then it is likewise properly integrable from a to b, and hence, by theorem IV,*

$$\int_a^b F(x)dx = \int_{a\,\Xi}^{b} F(x)dx.$$

This theorem is a corollary of the known

Lemma III. A necessary and sufficient condition for the proper integrability from a to b of a function $F(x)$ existent on the interval ab as a single valued limited function is this, that, with respect to a closed point set S of content zero, the function $F(x)$ be properly integrable from a' to b' where the interval $a'b'$ is any interval of ab containing no point s of the set S.

The theorem of lemma III was developed by PASCH* for the proper definite integral as defined by RIEMANN, and later independently by STOLZ† for the integral as defined by PEANO.‡ The two definitions are equivalent in content.

The lemma-theorem is a corollary of a théorem of DINI (1878: DINI-LÜROTH, § 185), which has close relations to the PEANO definition, and which is likewise given by STOLZ§ independently, but with reference to another theorem of DINI (DINI-LÜROTH, § 184), a particular case of the one here referred to.

VI. *If $F(x)$ is Ξ-integrable from a to b and if Ξ_0 is any subset of Ξ, such that for every interval-set I_0 enclosing Ξ_0 the function $F_{I_0}(x)$ is properly integrable from a to b, then $F(x)$ is Ξ_0-integrable from a to b, and*

---

* Mathematische Annalen, vol. 30 (1887), p. 147.

† Wiener Berichte, vol. 107 (1898), p. 736; *Grundzüge*, vol. 3 (1899), p. 272.

‡ *Lezioni di Analisi infinitesimale* (1893), § 103; Annali di Matematica, ser. 2, vol. 23 (1895), p. 157.

§ Monatshefte für Mathematik und Physik, vol. 7 (1896), p. 293.

$$\int_{a\,\Xi_0}^{b} F(x)\,dx = \int_{a\,\Xi}^{b} F(x)\,dx. \tag{13}$$

Remark. The set Ξ_0 includes the subset Z_{ab} lying on ab of the singular set Z of the function $F(x)$ (cf. § 1 4°), and it will appear by theorem VIIA of § 3 that Ξ_0 may be any such subset of Ξ.

We shall prove that

$$\left| \int_a^b F_{I_0}(x)\,dx - \int_{a\,\Xi}^{b} F(x)\,dx \right| < \epsilon \tag{14}$$

for every interval-set I_0 enclosing Ξ_0 $\begin{smallmatrix}\text{narrowly}\\ \text{broadly}\end{smallmatrix}$ and of length $D_{I_0} < \frac{1}{2}\delta_{\epsilon/2}$.

We extend I_0 by an interval-set J so that the interval-set I, $I = I_0 + J$, encloses Ξ $\begin{smallmatrix}\text{narrowly}\\ \text{broadly}\end{smallmatrix}$. Since Ξ has content zero and in accordance with the present hypothesis, in view of an obvious extension of lemma II from an interval ab, to an interval-set (in the present application, to the interval-set extending $I_{ab}^{\ b}$ to the interval ab), we are able to choose J so that $D_J < \frac{1}{2}\delta_{\epsilon/2}$ and

$$\left| \int_{J_a^b} F(x)\,dx \right| < \tfrac{1}{2}\epsilon. \tag{15}$$

Then $D_I < \delta_{\epsilon/2}$, and hence

$$\left| \int_a^b F_I(x)\,dx - \int_{a\,\Xi}^{b} F(x)\,dx \right| < \tfrac{1}{2}\epsilon. \tag{16}$$

Then the desired inequality (14) follows from (15), (16) since

$$\int_a^b F_I(x)\,dx + \int_{J_a^b} F(x)\,dx = \int_a^b F_{I_0}(x)\,dx.$$

§ 3.

FUNDAMENTAL PROPERTIES OF A FUNCTION $F(x)$ Ξ-INTEGRABLE FROM a TO b.

Fixed hypothesis. The following theorems concerning the function $F(x)$ involve the fixed hypothesis that the $\begin{smallmatrix}\text{narrow}\\ \text{broad}\end{smallmatrix}$ Ξ-integral

$$\int_{a\,\Xi}^{b} F(x)\,dx$$

exists.

1. *The function $F(x)$ is Ξ-integrable from b to a, and*

$$\int_{b\,\Xi}^{a} F(x)\,dx = -\int_{a\,\Xi}^{b} F(x)\,dx.$$

II. *The function $cF(x)$, where c is any constant, is Ξ-integrable from a to b, and*

$$\int_{a\,\Xi}^{b} cF(x)\,dx = c\int_{a\,\Xi}^{b} F(x)\,dx.$$

These two theorems follow immediately from the corresponding theorems concerning proper definite integrals together with the definition of the Ξ-integral.

III. *For every ϵ there is a δ_ϵ^1 such that the inequality:*

$$\left|\int_a^b F_{I_1}(x)\,dx - \int_a^b F_{I_2}(x)\,dx\right| < \epsilon,$$

holds for every pair of integral-sets $I_1 I_2$, each enclosing Ξ narrowly broadly and of length less than δ_ϵ^1.

This theorem is the necessary condition given in theorem II of § 2.

IV. *For every ϵ there is a δ_ϵ^2 such that*

$$\left|\int_{x_1}^{x_2} F_{I_1}(x)\,dx - \int_{x_1}^{x_2} F_{I_2}(x)\,dx\right| < \epsilon \tag{1}$$

for every two points $x_1 x_2$ of the interval ab and every pair of interval sets I_1, I_2 each enclosing Ξ narrowly broadly and of length less than δ_ϵ^2.

This theorem, whose proof will be given in connection with theorem X, of which it is a particular case, affirms that with respect to any two points x_1, x_2 of the interval ab the condition sufficient for the existence* of the Ξ-integral $\int_{x_1\,\Xi}^{x_2} F(x)\,dx$ is satisfied, and we obtain immediately theorems V and VI.

V. *The definite integral*

$$\int_{x_1\,\Xi}^{x_2} F(x)\,dx$$

exists where $x_1 x_2$ are any two points of the interval ab, and indeed uniformly† on the set of all such point-pairs $x_1 x_2$.

Corollary.‡ If the function $F(x)$ is Ξ-integrable on ab, then it is Ξ-integrable from a to b.

VI. *For every ϵ and every two points $x_1 x_2$ of the interval ab*

$$\left|\int_{x_1}^{x_2} F_I(x)\,dx - \int_{x_1\,\Xi}^{x_2} F(x)\,dx\right| \leqq \varepsilon \tag{2}$$

*At this point one observes that our definition of the integral $\int_{a\,\Xi}^{b} F(x)\,dx$ is more convenient than a definition applicable merely to cases in which the point-set Ξ lies on the interval ab.

† In that δ_ϵ^2 is independent of $x_1 x_2$.

‡ Cf. definition 3 and theorem III of § 2. The proof is indirect and quite analogous to the proof of the corresponding theorem as to proper definite integrals.

for every interval-set I enclosing Ξ $\overset{\text{narrowly}}{\text{broadly}}$ and of length less than δ_ϵ^2.

Remark. Theorem VI is useful especially in case $x_1 = a$, $x_2 = X$, when it establishes a relation on the X-interval ab between the various definite integral functions:

$$\int_{a\,\Xi}^{X} F(x)\,dx, \qquad \int_{a}^{X} F_I(x)\,dx,$$

where I denotes an interval-set enclosing Ξ $\overset{\text{narrowly}}{\text{broadly}}$.

VII A. *If $F(x)$ is Ξ-integrable from a to b and if Ξ_0 is any subset of Ξ which contains the subset Z_{ab} lying on ab of the singular set Z, then $F(x)$ is Ξ_0-integrable from a to b, and*

$$\int_{a\,\Xi_0}^{b} F(x)\,dx = \int_{a\,\Xi}^{b} F(x)\,dx = \int_{a\,Z}^{b} F(x)\,dx.$$

VII B. *If the function $F(x)$ is integrable from a to b with respect to each of two point-sets Ξ_1, Ξ_2 of content zero, then it is likewise with respect to their common subset Ξ_{12} and with respect to their aggregate set $\Xi_1 + \Xi_2$, and the four integrals are equal:*

$$\int_{a\,\Xi_1+\Xi_2}^{b} F(x)\,dx = \int_{a\,\Xi_1}^{b} F(x)\,dx = \int_{a\,\Xi_2}^{b} F(x)\,dx = \int_{a\,\Xi_{12}}^{b} F(x)\,dx.$$

The theorem VII A is a generalization of theorem VI of § 2, and, in the light of § 3 V and § 2 V, it is a corollary of that theorem, the function $F_{I_0}(x)$ for any interval-set I_0 enclosing Ξ_0 being on ab a single valued limited function of x, as one proves from the definition of the singular set Z (§ 1, 4°) indirectly by the usual interval-halving process.

As to the theorem VII B, the sets Ξ_1, Ξ_2, Ξ_{12} have the set Z of essential singularities as a common subset. Hence by VII A the Ξ_{12}-integral exists and we have:

$$\int_{a\,\Xi_1}^{b} F(x)\,dx = \int_{a\,\Xi_2}^{b} F(x)\,dx = \int_{a\,\Xi_{12}}^{b} F(x)\,dx.$$

Denoting for the moment by δ_ϵ the least of the δ_ϵ's related to these three integrals, we consider an interval-set I enclosing $\Xi_1 + \Xi_2$ $\overset{\text{narrowly}}{\text{broadly}}$ and of length $D_I < \delta_{\epsilon/3}$. Then, setting in a definite way:

$$I = I_{12} + I_1 + I_2,$$

where I_{12} encloses Ξ_{12} narrowly, and $I_{12} + I_1$ encloses Ξ_1 $\overset{\text{narrowly}}{\text{broadly}}$, and $I_{12} + I_2$ encloses Ξ_2 $\overset{\text{narrowly}}{\text{broadly}}$, we have obviously, in view of the equality of the three integrals, the three inequalities:

$$\left| \int_a^b F_J(x)\,dx - \int_{a\,\Xi_{12}}^{b} F(x)\,dx \right| < \tfrac{1}{3}\epsilon \quad (J = I_{12},\ I_{12} + I_1,\ I_{12} + I_2),$$

from which, in view of the fact that $F(x)$ is properly integrable on I_1 and I_2, we deduce readily the inequality:

$$\left| \int_a^b F_I(x)\,dx - \int_{a\,\Xi_{12}}^{b} F(x)\,dx \right| < \epsilon,$$

needed to complete the proof of the theorem.

VIII. *The relation:*

$$\int_{x_1\,\Xi}^{x_2} F(x)\,dx + \int_{x_2\,\Xi}^{x_3} F(x)\,dx = \int_{x_1\,\Xi}^{x_3} F(x)\,dx, \tag{3}$$

holds for every three points $x_1 x_2 x_3$ of the interval ab.

This follows easily from the corresponding theorem concerning proper definite integrals. We have for every interval-set I enclosing Ξ $^{\text{narrowly}}_{\text{broadly}}$

$$\int_{x_1}^{x_2} F_I(x)\,dx + \int_{x_2}^{x_3} F_I(x)\,dx = \int_{x_1}^{x_3} F_I(x)\,dx. \tag{4}$$

For an interval-set I of length $D_I < \delta_\epsilon^2$ the two sides of the equality (4) differ respectively from the two corresponding sides of (3) by less than 2ϵ. Hence the two sides of (3) differ by less than 4ϵ and so are indeed equal.

IX. *The definite integral function:*

$$\phi(X) = \int_{a\,\Xi}^{X} F(x)\,dx,$$

is a continuous function of the variable upper limit X on the X-interval ab.

Introducing with respect to an ϵ an interval-set I enclosing Ξ narrowly and of length $D_I < \delta_{\epsilon/2}^2$, we have by the use of theorems VIII and VI the relation:

$$|\phi(X_1) - \phi(X_2)| < \epsilon, \tag{5}$$

where X_1 and X_2 are any two points of the X-interval ab such that

$$|X_1 - X_2| < \delta'_{\epsilon/2},$$

the $\delta'_{\frac{1}{2}\epsilon}$ being the number related to the properly integrable function $F_I(x)$ quâ $f(x)$ by lemma II of § 2. And this relation establishes the uniform continuity of $\phi(X)$ on the X-interval ab. This theorem is generalized in theorems XIII′ and XIII″.

REMARK. Up to this point of the present development the theories of the narrow and of the broad Ξ-integrals are equivalent. From this point on they diverge; I secure for the narrow Ξ-integrals conclusions analogous to the conclusions for the broad Ξ-integrals, by inserting an additional hypothesis in the first case of every theorem.

X. *For every ϵ there is a δ_ϵ^2 such that for every interval-set J lying on the interval ab*

$$\left| \int_J F_{I_1}(x)\,dx - \int_J F_{I_2}(x)\,dx \right| < \epsilon \tag{6}$$

for every pair $I_1 I_2$ of interval-sets satisfying the following conditions:

(1) *Each interval-set $I_1 I_2$ encloses Ξ narrowly broadly;*

(2) *Each length $D_{I_1} D_{I_2}$ is less than δ_ϵ^2;*

and, moreover, in the first case, whenever J consists of more than a single interval,

(3) *Every interval i of I_1 or I_2 which encloses an interval l joining two consecutive intervals of J has with respect to every such enclosed interval l the following property: the interval i contains either on the interval l a point ξ or on each of the adjoining intervals of J a point ξ.*

Remark. In case J consists of a single interval, theorem X becomes theorem IV; and in case J consists of more than a single interval, the third condition is fulfilled in particular if there is imposed on J an obvious restriction, and likewise, if there are imposed on I_1 and I_2 obvious restrictions, which have reference to the interval-set J in question. In the theorem as here stated, the uniformity with respect to the set of all interval-sets I_1, I_2, J of the nature specified is especially important for the sequel.

We use the determination

$$\delta_\epsilon^2 = \tfrac{1}{2}\delta_\epsilon^1, \tag{7}$$

and have from theorem III the inequality:

$$\left| \int_a^b F_{I_1}(x)\,dx - \int_a^b F_{I_2}(x)\,dx \right| < \epsilon, \tag{8}$$

valid for every ϵ and for every pair of interval-sets $I_1 I_2$ each enclosing Ξ narrowly broadly and of length less than $2\delta_\epsilon^2$.

For the purposes of an indirect proof of theorem X, we suppose that the inequality:

$$\left| \int_J F_{I_1}(x)\,dx - \int_J F_{I_2}(x)\,dx \right| \geqq \epsilon, \tag{9}$$

holds for a certain ϵ, a certain interval-set J lying on the interval ab, and a certain pair of interval-sets $I_1 I_2$ each enclosing Ξ narrowly broadly and of length less than δ_ϵ^2, and, in the first case, satisfying the condition (3) of X, and proceed to exhibit for this ϵ two interval-sets $I_3 I_4$ each enclosing Ξ narrowly broadly and of length less than $2\delta_\epsilon^2$ for which, in contradiction with (8), the inequality:

$$\left| \int_a^b F_{I_3}(x)\,dx - \int_a^b F_{I_4}(x)\,dx \right| \geqq \epsilon, \tag{10}$$

has validity.

It is easy to see how the inequality (9) may be transformed into an inequality of the general form (10). For denoting by K the interval-set obtained by excising J from the interval ab, we have for every pair of sets $I_3 I_4$ enclosing Ξ, the equation:

$$(11)\quad \int_a^b F_{I_3}(x)\,dx - \int_a^b F_{I_4}(x)\,dx = \begin{cases} \left(\int_J F_{I_3}(x)\,dx - \int_J F_{I_4}(x)\,dx\right) \\ + \left(\int_K F_{I_3}(x)\,dx - \int_K F_{I_4}(x)\,dx\right), \end{cases}$$

since the functions $F_{I_3}(x)$, $F_{I_4}(x)$ are properly integrable from a to b. Hence from (9) we obtain the relation (10) for every pair of sets $I_3 I_4$ each enclosing Ξ and such that

$$(12)\quad \begin{aligned} F_{I_3}(x) &= F_{I_1}(x), \qquad F_{I_4}(x) = F_{I_2}(x) && (x \text{ within } J), \\ F_{I_3}(x) &= F_{I_4}(x) && (x \text{ within } K), \end{aligned}$$

the various proper definite integrals having values independent of the values of the respective integrand functions at the extremities of the intervals of J and K.

Thus we seek to find two sets $I_3 I_4$ each enclosing Ξ $\begin{matrix}\text{narrowly}\\ \text{broadly}\end{matrix}$ and of length less than $2\delta_\epsilon^2$, and, moreover, such that they agree on J with $I_1 I_2$ respectively and on K with each other.

We denote by L the complete x-axis with the omission of the inner points of J. Thus L contains K. And we denote by $I_J I_L$ respectively the J-section and the L-section of an interval-set I, so that

$$(13)\quad I = I_J + I_L.$$

With respect to J and L an interval i may lie on J or L and so be itself an interval i_J or i_L. Otherwise the interval i is by J and L partitioned into a sequence of two or more parts i_J, i_L, two adjoining parts being a part i_J and a part i_L. The two extreme parts may or may not be undivided intervals j or l. The intermediate parts are undivided intervals j or l.

We must for an interval-set I enclosing Ξ $\begin{matrix}\text{narrowly}\\ \text{broadly}\end{matrix}$ distribute the intervals i_L of I_L into two complementary classes.

An interval i_L is an interval i_L'' if it contains no point ξ and is either an extreme part of the interval i of I from which it is derived or an intermediate part each of whose adjoining parts i_J contains a point ξ. An interval i_L is an interval i_L' if it either contains a point ξ or is of its interval i an intermediate part one (at least) of whose adjoining intervals i_J contains no point ξ.

Thus we have two interval-sets $I_L' I_L''$ such that $I_L = I_L' + I_L''$.

In case I encloses Ξ narrowly we notice two facts which are of use in the immediate sequel. An interval i which lies on J or L is an interval i_J or i'_L. An interval i undergoing partition with respect to J and L contains at least one part i_J and the sum of its parts i_J, i'_L is an interval-set every interval of which contains a point ξ. These intervals of i are the intervals arising by the excision from i of its various parts i''_L. The interval i itself contains a point ξ. If it contains no part i''_L it itself is the interval in question. If, however, the interval i contains one or more parts i''_L, we consider any such part i''_L. If i''_L is an extreme part the adjoining part i_J contains a point ξ or has in progressive consecution an adjoining part i'_L, since an adjoining part i''_L would be an extreme part, and the interval i consisting of these three parts would have no point ξ. If this third part has no point ξ, there is a fourth part, i_J. If this part i_J has no point ξ, there is a fifth part i'_L. Thus the parts i_J, i'_L enter in alternation and the interval consisting of this complete and unbroken sequence of parts i_J, i'_L contains a point ξ. If on the other hand i''_L is an intermediate part each adjoining part i_J contains a point ξ.

In these notations we take $I_3 I_4$ as follows:*

$$I_3 = I_{1J} + I'_{1L} + I'_{2L}, \quad I_4 = I_{2J} + I'_{1L} + I'_{2L}. \tag{14}$$

And these sets satisfy the prescribed conditions:

(1) Each set is of length less than $2\delta_\epsilon^2$;

(2) $$I_{3J} = I_{1J}, \quad I_{4J} = I_{2J};$$

(3) $$I_{3L} = I_{4L} = I'_{1L} + I'_{2L};$$

(4) Each set encloses Ξ $\begin{matrix}\text{narrowly}\\ \text{broadly}\end{matrix}$.

We need to prove that I_3 and so I_4 encloses Ξ $\begin{matrix}\text{narrowly}\\ \text{broadly}\end{matrix}$. The set Ξ is by hypothesis closed.

We consider first any particular point ξ. If ξ lies within J or L it is enclosed by an interval of I_{1J} or I'_{1L}, and so by an interval of I_3. If, however, ξ is common to J and L it is an extremity of an interval of each set; the point ξ lies within a certain interval of I_1, and it is the point of junction of two intervals of $I_{1J} I'_{1L}$ respectively, and in $I_3 = I_{1J} + I'_{1L} + I'_{2L}$ these two intervals perhaps with others form an interval i_3 enclosing the point ξ. Thus the interval-set I_3 encloses Ξ at least broadly and theorem X is proved for the second case.

For the first case we consider any particular interval i_3 of I_3 and now need to prove merely that it contains a point ξ. The interval i_3 arises by the union of certain intervals of $I_{1J} I'_{1L} I_{2L}$. In the introduction of the notation I'_L

* In the second case, one may determine suitable interval-sets I_3, I_4 more simply as follows:

$$I_3 = I_{1J} + I''', \quad I_4 = I_{2J} + I''',$$

where I''' is the interval-set common to I_{1L} and I_{2L}.

for an interval-set $I = I(\Xi)$ it was noted that every interval i_J or i''_L forms part of a sequence of intervals $i_J i''_L$ constituting an interval containing a point ξ. Thus i_3 contains a point ξ or it consists of a single interval i''_{2L} which contains no point ξ. But this latter case is impossible. For an interval i''_{2L} containing no point ξ is an intermediate part of an interval i_2 one of whose adjoining parts i_{2J} contains no point ξ. But this is impossible, by virtue of the third hypothesis of theorem X.

Thus, indeed, every interval i_3 contains a point ξ. And now theorem X is proved also for the first case, and, hence, completely.

XI. *For every interval-set J lying on the interval ab there exists a constant, in notation,*

$$\int_{J\Xi} F(x)\,dx, \tag{15}$$

the integral on the interval-set J with respect to the point-set Ξ of the function $F(x)$, such that for every ϵ the relation:

$$\left| \int_J F_I(x)\,dx - \int_{J\Xi} F(x)\,dx \right| \leqq \epsilon, \tag{16}$$

holds for every interval-set I satisfying the following conditions:

(1) *The set I encloses Ξ narrowly / broadly;*

(2) *The length D_I is less than δ_ϵ^2;*

and, moreover, in the first case, whenever J consists of more than a single interval,

(3) *Every interval i of I which encloses an interval l joining two consecutive intervals of J has with respect to its every such enclosed interval l the following property: the interval i contains either on the interval l a point ξ or on each of the adjoining intervals of J a point ξ.*

This theorem follows directly from the preceding theorem by the usual limit-considerations, wherein it needs to be noticed that interval-sets $I_1 I_2$ satisfying the conditions 1, 2, 3 of that theorem actually exist for every ϵ and interval-set J of the interval ab.

Remark. If J is a single interval $a'b'(a' < b')$ then obviously

$$\int_{J\Xi} F(x)\,dx = \int_{a'\Xi}^{b'} F(x)\,dx. \tag{17}$$

This remark is generalized in the following theorem:

XII. *For every interval-set J of the interval ab,*

$$\int_{J\Xi} F(x)\,dx = \sum_{l=1}^{n} \int_{a_l\Xi}^{b_l} F(x)\,dx, \tag{18}$$

where the intervals $a_l b_l (a_l < b_l;\ l = 1, \cdots, n)$ are the intervals of J.

For every interval-set I enclosing Ξ we have, by the definition (11) of § 2,

$$\int_J F_I(x)\,dx = \sum_{l=1}^{n} \int_{a_l}^{b_l} F_I(x)\,dx. \tag{19}$$

We introduce a set I which satisfies with respect to J and so with respect to every interval $a_l b_l$ of J the conditions of theorem XI for $\epsilon = \epsilon'/(n+1)$ where ϵ' is any particular positive constant. Then by theorem XI we have

$$\left| \int_J F_I(x)\,dx - \int_{J\Xi} F(x)\,dx \right| \leqq \frac{\epsilon'}{n+1}, \tag{20}$$

$$\left| \int_{a_l}^{b_l} F_I(x)\,dx - \int_{a_l \Xi}^{b_l} F(x)\,dx \right| \leqq \frac{\epsilon'}{n+1} \qquad (l = 1, 2, \cdots, n), \tag{21}$$

and then from (19), (20), (21) we have

$$\left| \int_{J\Xi} F(x)\,dx - \sum_{l=1}^{n} \int_{a_l \Xi}^{b_l} F(x)\,dx \right| \leqq \epsilon', \tag{22}$$

from which the theorem follows at once.

Remark. Theorem XI is to be understood as applying also to the case in which the symbol J denotes a non-existent interval-set, the constant in question having the value 0. Then in view of theorem XII it is evident that with respect to an interval-set J lying on the interval ab and a function $F(x)$ Ξ-integrable from a to b a definition of the symbol $\int_{J\Xi} F(x)\,dx$ might be given similar to that given in § 2 for the symbol $\int_J f(x)\,dx$ with respect to a function $f(x)$ properly integrable from a to b.

XIII′. *For every ϵ there exists a δ_ϵ^3 such that*

$$\left| \int_{J\Xi} F(x)\,dx \right| < \epsilon \tag{23}$$

for every interval-set J satisfying the following conditions:

(1) *The interval-set J lies on the interval ab;*

(2) *The length D_J is less than δ_ϵ^3;*

and, moreover, in the first case, whenever J consists of more than a single interval,

(3′) *The interval-set J has with respect to every interval l joining two consecutive intervals of J the following property: either the interval l contains a point ξ or each of the adjoining intervals of J contains a point ξ.*

XIII″. *In the first case: For every two positive numbers $\epsilon\epsilon'$ there exists a $\delta_{\epsilon\epsilon'}^3$ such that*

$$\left| \int_{J\Xi} F(x)\,dx \right| < \epsilon \tag{23}$$

for every interval-set J satisfying the following conditions:

(1) *The interval-set J lies on the interval ab;*

(2) *The length D_J is less than $\delta^3_{\epsilon\epsilon'}$;*

(3″) *Whenever J consists of more than a single interval, and has an interval l joining two consecutive intervals of J, and lying within the ϵ'-neighborhood of a point ξ, it has with respect to every such interval l the following property: either the interval l contains a point ξ or each of the adjoining intervals of J contains a point ξ.*

Note 1. The condition 3′ of XIII′ is satisfied in particular by interval-sets J satisfying the condition:

(3‴) *Every interval of J contains a point ξ.*

Note 2. For an interval-set I enclosing Ξ $^{\text{narrowly}}_{\text{broadly}}$ and of length less than δ^3_ϵ the inequality

$$\left| \int_{I_a^b \Xi} F(x)\,dx \right| < \epsilon$$

holds. This is a particular case of XIII′. With δ^3_ϵ replaced by δ_ϵ, it follows from the definition of the Ξ-integral from a to b, in view of the relation:

$$\int_{I_a^b \Xi} F(x)\,dx = \int_{a\Xi}^{b} F(x)\,dx - \int_a^b F_I(x)\,dx,$$

derived from theorem XII, together with § 2 IV.

Theorem XIII is for functions Ξ-integrable a generalization* of the second lemma of § 2 for functions properly integrable and it is proved by means of that lemma and theorem XI.

We take a particular interval-set I enclosing Ξ $^{\text{narrowly}}_{\text{broadly}}$ and of length D_I less than $\delta^2_{\epsilon/2}$, and moreover, in connection with theorem XIII″, of intervals each of length at most ϵ'. The function $F_I(x)$ is properly integrable from a to b.

An effective number δ^3_ϵ, or, for the second theorem of the first case, $\delta^3_{\epsilon\epsilon'}$, is the smaller of certain numbers: viz., the number $\delta'_{\epsilon/2}$ connected with this function $F_I(x)$ quâ $f(x)$ by the second lemma of § 2, and if I_a^b consists of more than a single interval, the various lengths of the intervals joining two consecutive intervals of I_a^b. Thus an interval j of an interval-set J of length $D_J < \delta^3_\epsilon$ or $\delta^3_{\epsilon\epsilon'}$ has points in common with at most one interval i of I.

This number δ^3_ϵ or $\delta^3_{\epsilon\epsilon'}$ is indeed effective. For the inequality (23) follows at once from the simultaneous inequalities:

$$\left| \int_J F_I(x)\,dx \right| < \tfrac{1}{2}\epsilon, \tag{24}$$

* The inequality (23) of XIII holds if on the right ϵ be replaced by 2ϵ and on the left the absolute value of the Ξ-integral over the interval-set J be replaced by the sum of the absolute values of the integrals over the constituent intervals of J. This is seen by considering separately the intervals yielding positive integrals and those yielding negative intergals. Cf. JORDAN, loc. cit., p. 226, ll. 9, 10; pp. 224–5.

$$\left| \int_J F_I(x)\,dx - \int_{J\,\Xi} F(x)\,dx \right| \leqq \tfrac{1}{2}\epsilon, \tag{25}$$

where J is an interval-set satisfying the conditions of theorem XIII. The inequality (24) is by virtue of the second lemma of § 2, since $D_J < \delta'_{\epsilon\,2}$. The inequality (25) is by virtue of theorem XI.

And theorem XI is, indeed, applicable. For J lies on the interval ab, and I encloses Ξ $\begin{matrix}\text{narrowly}\\ \text{broadly}\end{matrix}$ and is of length D_I less than $\delta^2_{\epsilon\,2}$. Thus the double theorem XIII is proved, in the second case, and also in the first case, if J consists of a single interval.

In the first case, if J consists of more than a single interval, we have further to prove that condition (3) of the hypothesis of theorem XI is satisfied. We consider the interval-sets J and I with respect to an interval l joining two consecutive intervals j', j'' of J. If l contains a point ξ or if j' and j'' each contain a point ξ, then an interval i of I enclosing l, if one there be, has with respect to l the property of condition (3) of theorem XI, that is, the interval i contains either on the interval l a point ξ or on each interval j', j'' a point ξ. For each point ξ lies in some interval of I, and since j' and j'' intersect i neither intersects a second interval of I. If, however, l contains no point ξ and either j' or j'' contains no point ξ, then no interval i encloses l; for we have now to do with XIII″, and by hypothesis (3″), the interval l lies within the ϵ'-neighborhood of no point ξ, while every interval i contains a point ξ and is of length at most ϵ'.

XIV. *The function $F_I(x)$ is properly integrable from a to b and the inequality:*

$$\left| \int_a^b F_I(x)\,dx - \int_{a\,\Xi}^{b} F(x)\,dx \right| < \epsilon, \tag{26}$$

holds for every interval-set I whose length D_I is less than δ^3_ϵ or $\delta^3_{\epsilon\epsilon'}$, and moreover, in the first case, whose interval-set I^b_a, quâ interval-set J satisfies the condition* (3′) or (3″) of theorem* XIII, *provided moreover (in each case) that the interval-set I contains the singular point-set Z of the function $F(x)$ in such a way that if a singular point ζ is an extremity of an interval i of I it is singular merely towards the interior of i.*

Note. This theorem is, in accordance with note 2 of theorem XIII, *of the nature of an extension of the definitional property of the function $F(x)$ Ξ-integrable from a to b.*—In particular, we see that *the relation* (26) *holds for every interval-set I of length $D_I < \delta^3_\epsilon$ whose corresponding function $F_I(x)$ is properly integrable from a to b, which contains Ξ, and, in the first case, whose every*

* The $\delta^3_{\epsilon\epsilon'}$ enters only in the first case and then only if the hypothesis imposes (merely) the condition (3″) of theorem XIII″.

interval contains a point ξ of Ξ. Such an interval-set I obviously may perhaps not enclose Ξ, for an interval i may have a limit-point ξ' of Ξ as an extremity.

We proceed to prove theorem XIV. The singular point-set Z was defined in §1 4°. From the final condition as to the interval-set I it follows by the usual process of indirect proof that the function $F_I(x)$ is on the interval ab everywhere defined, single valued, and limited. It is moreover Ξ-integrable from a to b, by virtue of theorems III and IV of §2, and V and VIII of §3. Thus by theorem V of §2 it is properly integrable from a to b. Further by these theorems and theorem XII we have the equality:

$$(27) \qquad \int_a^b F_I(x)\,dx = \int_{a\,\Xi}^{b} F_I(x)\,dx = \int_{a\,\Xi}^{b} F(x)\,dx - \int_{I_a^b\,\Xi} F(x)\,dx,$$

from which the theorem follows by virtue of theorem XIII.

XV. *If the function $F(x)$ is integrable from a to b with respect to a closed point-set Ξ of content zero, and if* H *is any point-set of content zero, such that $\Xi + \mathrm{H}$ is closed, and which, moreover, in the first case, satisfies the condition specified below, then the function $F(x)$ is $(\Xi + \mathrm{H})$-integrable from a to b, and*

$$(28) \qquad \int_{a\,\Xi+\mathrm{H}}^{b} F(x)\,dx = \int_{a\,\Xi}^{b} F(x)\,dx.$$

The set H *in the first case has the following property: there exist on the interval ab at most a finite number of intervals $a'b'$ enclosing no point ξ and no point η and having for extremities points η which are not points ξ.*

In the second case this theorem is evident, since the interval-sets I enclosing $\Xi + \mathrm{H}$ broadly are amongst those enclosing Ξ broadly.

We proceed to prove the theorem in the first case. An interval-set K enclosing $\Xi + \mathrm{H}$ narrowly is with respect to Ξ separated into an interval-set I enclosing Ξ narrowly and a complementary interval set J perhaps not existent.

With respect to a number ϵ we shall prove that

$$(29) \qquad \left| \int_a^b F_K(x)\,dx - \int_{a\,\Xi}^{b} F(x)\,dx \right| < \epsilon$$

for every such interval-set K of length $D_K < \delta_\epsilon^1$, where δ_ϵ^1 is the least of the three numbers $\delta_{\epsilon/2}$, ϵ', and $\delta^1_{\epsilon/2\,\epsilon'}$, where ϵ' is a certain positive number to be specified in the sequel.

Indeed the inequality (29) follows at once from the fact that the relations:

$$(30) \qquad \int_a^b F_K(x)\,dx = \int_a^b F_I(x)\,dx - \int_{J_a^b\,\Xi} F(x)\,dx,$$

$$\left| \int_a^b F_I(x)\,dx - \int_{a\,\Xi}^{b} F(x)\,dx \right| < \tfrac{1}{2}\,\epsilon, \tag{31}$$

$$\left| \int_{J_a^b\,\Xi} F(x)\,dx \right| < \tfrac{1}{2}\,\epsilon, \tag{32}$$

hold simultaneously. Of these relations the first and second are clear, and the third follows from theorem XIII'', in connection with the stipulations $D_{J_a^b} \leqq D_K < \delta_\epsilon^1 < \epsilon'$, $\delta_{\epsilon 2\epsilon'}^3$. In order to secure this application of theorem XIII'' we are to determine the positive number ϵ' so that every interval l joining two consecutive intervals of J_a^b either contains a point ξ or lies within the ϵ'-neighborhood of no point ξ. This determination is made by means of the final condition of the theorem.

We denote by H* the set of points $\eta = \eta^*$ which lie on the interval ab and which are not points ξ; by H_* the set obtained by adding to H* the points a and b; and for brevity by J_0 the set J_a^b.

An interval j of J contains a point η and no point ξ. An interval j_0' of J_0 contains a point η_* and no point ξ. Two consecutive intervals j_0', j_0'' of J determine an interval $a'b'$ where a' is that point η_* of j_0' nearest to j_0'', and where similarly b' is that point η_* of j_0'' nearest to j_0'. These points $a'b'$ are definite points, for the sets H_* on the two intervals $j_0' j_0''$ are closed sets. This interval $a'b'$ encloses no ξ and no η,—except perhaps within the interval l joining j_0' and j_0''. Now if l encloses no point ξ it encloses no interval i of I, and hence, since $j_0' j_0''$ are consecutive intervals of J_0, it encloses no point η. Thus the two consecutive intervals $j_0' j_0''$ of J_0 determine a joining-interval l containing a point ξ or else an interval $a'b'$ enclosing no point ξ and no point η and having for extremities points η_*.

Now a enters as an η_*-extremity of such an interval at most once, and likewise b enters at most once. And by the present hypothesis at most a finite number of such intervals $a'b'$ with η^*-extremities exist.

We denote by $3\epsilon'$ the least of the lengths of the various intervals $a'b'$ with η_*-extremities, setting $3\epsilon' = 1$, if no such interval exists. Thus ϵ' is a definite positive number.

Then, since $D_K < \epsilon'$, the intervals $j_0' j_0''$ are each of length less than ϵ', while their interval $a'b'$, whenever it contains no point ξ, is of length at least $3\epsilon'$ and hence their joining-interval l, whenever it contains no point ξ, is of length greater than ϵ' and so lies within the ϵ'-neighborhood of no point ξ.—Thus theorem XIII'' is available for the completion of the proof of theorem XVII.

XVI. *In the first case: If the function $F(x)$ is Ξ-integrable from a to b, and hence* * *Z-integrable, if however it is not* (Z + H)*-integrable, and hence* *

* Theorem VIIA.

not $(\Xi + \mathrm{H})$*-integrable, where* H *is a closed point-set of content zero, then there is at least one point* $\zeta = \zeta_0$ *of* Z *on* ab *at which* $F(x)$ *fails either of progressive or of regressive* $(\mathrm{Z} + \mathrm{H})$*- and so of* $(\Xi + \mathrm{H})$*-integrability. Hence, by theorem* XV, *in every progressive or in every regressive neighborhood of such a point* ζ_0 *there is an infinitude of intervals* $a'b'$ *enclosing no point* ξ *and no point* η *and having for extremities points* η *which are not points* ξ.

For if there were no such point ζ_0, then $F(x)$ would be $(\mathrm{Z} + \mathrm{H})$-integrable on ab and hence from a to b. We need to recall the statements of 4° of § 1, the definitions 2 and 3 and theorem IV of § 2, and the corollary of theorem V of § 3.

Remark. According to theorem XV the hypotheses of theorem XVI imply that the Ξ-integral of $F(x)$ is essentially narrow (that is, not broad). It has not been proved however that, in this case, there exists a point-set H for which the narrow $(\Xi + \mathrm{H})$-integral of $F(x)$ is non-existent.

XVII. *If with respect to two closed point-sets* Ξ, H *of content zero two functions* $F(x)$, $G(x)$ *are respectively integrable from* a *to* b, *and if, in the first case,* Ξ, H *are so related (for instance, as indicated in theorem* XV) *that each function is* $(\Xi + \mathrm{H})$*-integrable from* a *to* b, *then* $F(x) + G(x)$ *is* $(\Xi + \mathrm{H})$*-integrable from* a *to* b, *and*

$$\int_{a\,\Xi}^{b} F(x)\,dx + \int_{a\,\mathrm{H}}^{b} G(x)\,dx = \int_{a\,\Xi+\mathrm{H}}^{b} \big(F(x) + G(x)\big)\,dx.$$

This theorem is easily seen to be true.

§ 4.

Concerning the absolutely convergent $\overset{\text{narrow}}{\text{broad}}$ Ξ-integrals.

Definition. The $\overset{\text{narrowly}}{\text{broadly}}$ existent integral $\int_{a\,\Xi}^{b} F(x)\,dx$ is said to converge or exist (as a limit) *absolutely* in case the integral $\int_{a\,\Xi}^{b} |F(x)|\,dx$ exists $\overset{\text{narrowly}}{\text{broadly}}$.

Making suitable use of the references to the work of Jordan and Stolz given in the introduction, the reader will readily construct the proofs of the following theorems.

I. *The upper limits*

$$\underset{I\,I\,(\Xi)}{\mathrm{L}} \int_a^b F_I(x)\,dx, \quad \underset{I\,I\,(\Xi)}{\mathrm{L}} \int_a^b F_I(x)\,dx$$

with respect to a non-negative function $F(x)$ *on the set of all the interval-sets* I *enclosing* Ξ *respectively narrowly or broadly are equal. If they are finite, the* Ξ*-integral*

$$\int_{a\,\Xi}^{b} F(x)\,dx$$

exists broadly and is equal to the common value of the upper limits. Conversely, if the Ξ*-integral exists (even) narrowly, then the upper limits are finite, and the* Ξ*-integral exists broadly.*

II. *If the integral*

$$\int_{a\,\Xi}^{b} |F(x)|\,dx$$

exists (if narrowly, then broadly), then the integral

$$\int_{a\,\Xi}^{b} F(x)\,dx$$

exists broadly, and hence as a broad integral it converges absolutely. Conversely, if the latter integral exists absolutely (even) narrowly, it exists also absolutely broadly.

III. *The broad and absolutely convergent* Ξ*-integrals are in fact identical; hence, the essentially narrow and the non-absolutely convergent* Ξ*-integrals are in fact identical.*

§ 5.*

Concerning the general or narrow Ξ-integrals

$$\int_{a\,(\Xi)}^{b} F(x)\,dx.$$

1°. The following determination of the general point-set Ξ lying on ab closed and of content zero is well known.

Every finite or numerably infinite set H of intervals $h_\nu = a_\nu b_\nu (\nu = 1, 2, \cdots, n$ or $\nu = 1, 2, 3, \cdots)$ lying everywhere densely on ab and no two having a common inner point determines a closed nowhere dense point-set Ξ; and, conversely, every closed nowhere dense point-set Ξ of ab is so determinable by its "point-free" intervals h_ν. Every inner point x of ab not a point ξ lies lies within a definite interval h_ν. The points ξ together with a and b are the extremities $a_\nu b_\nu$ of the intervals h_ν and the limit-points of such extremities.— For later use I introduce subintervals $h'_\nu = a'_\nu b'_\nu$ of the respective intervals $h_\nu = a_\nu b_\nu$, with the understanding that $a'_\nu = a_\nu$ if $a_\nu = a =$ not-ξ; $b'_\nu = b_\nu$ if $b_\nu = b =$ not-ξ; and otherwise $a_\nu < a'_\nu < b'_\nu < b_\nu$.

The content of the set Ξ so determined is $|a-b| - \sum_\nu |a_\nu - b_\nu|$, so that for our purposes

$$\sum_\nu |a_\nu - b_\nu| = |a-b|. \tag{1}$$

2°. We consider any such set Ξ and have the following important theorem:

* Addition of July 1, 1901.

THEOREM. *For the existence of the narrow* Ξ*-integral*

$$\int_{a(\Xi)}^{b} F(x)\,dx \tag{2}$$

a necessary and sufficient (twofold) condition is this, that

(α) *The narrow* Ξ_ν*-integrals*

$$\int_{a_\nu(\Xi_\nu)}^{b_\nu} F(x)\,dx \qquad (\Xi_\nu = (a_\nu, b_\nu)); \tag{3}$$

exist;

and, if Ξ *contains an infinitude of points,*

(β) *The infinite series*

$$\sum_{\nu=1}^{\infty} O_\nu \tag{4}$$

converges, where for every O_ν *denotes the oscillation on the* X*-interval* $a_\nu b_\nu$ *of the continuous definite integral function*

$$J_\nu(X) = \int_{a_\nu(\Xi_\nu)}^{X} F(x)\,dx. \tag{5}$$

This sufficient twofold condition being satisfied, the integral (2) *has the value*

$$\int_{a(\Xi)}^{b} F(x)\,dx = \sum_\nu \int_{a_\nu(\Xi_\nu)}^{b_\nu} F(x)\,dx. \tag{6}$$

As to the final statement, it is obvious that the terms of the series on the right of (6) are by (α) definite numbers, and that the series (6) if infinite is convergent by (β).

The twofold condition is necessary. In fact, § 3 V gives (α) at once, while (β) follows indirectly * easily from § 3 XIII′. For the oscillation O_ν may be expressed in the form

$$O_\nu = \int_{X_\nu''(\Xi_\nu)}^{X_\nu'} F(x)\,dx = \int_{X_\nu''(\Xi)}^{X_\nu'} F(x)\,dx, \tag{7}$$

where $X_\nu' X_\nu''$ are points of $a_\nu b_\nu$ at which $J_\nu(X)$ has its maximum and minimum values respectively. Now we suppose that the infinite series (4) diverges, and we consider any two positive numbers $\epsilon\,\delta$. In view of (1) there is an integer ν_δ such that

$$\sum_{\nu > \nu_\delta} |a_\nu - b_\nu| < \delta. \tag{8}$$

* The direct proof is equally simple.

Then in connection with the divergent series

$$\sum_{\nu>\nu_\delta} O_\nu \tag{9}$$

there is an integer $\nu_{\delta,\epsilon} > \nu_\delta$ such that

$$\sum_{\nu=\nu_\delta-1}^{\nu=\nu_{\delta\epsilon}} O_\nu > 2\epsilon. \tag{10}$$

The terms O_ν of (10) separate into two sets according as $X'_\nu > X''_\nu$ or $X'_\nu < X''_\nu$, and thereby determine two interval-sets J_1, J_2 of intervals $X''_\nu X'_\nu$, $X'_\nu X''_\nu$, each lying on ab and of length less than δ, and each having the property that every interval l joining two of its intervals contains a point ξ, and furthermore either J_1 or J_2 quâ J having the property that

$$\left| \int_{J(\Xi)} F(x)\,dx \right| > \epsilon. \tag{11}$$

And this result is in contradiction with §3 XIII′.

It remains to prove that the twofold condition is sufficient. For a finite set Ξ this follows at once from §2 III. We consider* then an infinite set Ξ, so that $\nu = 1, 2, \cdots$. For every ϵ a δ_ϵ must be exhibited such that for every interval-set I enclosing Ξ narrowly and of length $D_I < \delta_\epsilon$ we have

$$\left| \int_a^b F_I(x)\,dx - \sum_\nu \int_{a_\nu(\Xi_\nu)}^{b_\nu} F(x)\,dx \right| < \epsilon. \tag{12}$$

We denote by J the complementary interval-set on ab of I_a^b; J consists then of say m intervals $a'_{\nu'}, b'_{\nu'}$ $(\nu' = \nu_l;\ l = 1, 2, \cdots, m)$; thus we have

$$\int_a^b F_I(x)\,dx = \int_J F(x)\,dx = \sum_{\nu'} \int_{a'_{\nu'}}^{b'_{\nu'}} F(x)\,dx. \tag{13}$$

Here we have denoted certain m indices ν by ν'. Denoting the remaining indices ν by ν'', we have the desired inequality (12) in the more convenient form

$$\left| \sum_{\nu'} \left(\int_{a_{\nu'}(\Xi_{\nu'})}^{b_{\nu'}} - \int_{a'_{\nu'}}^{b'_{\nu'}} \right) F(x)\,dx + \sum_{\nu''} \int_{a_{\nu''}(\Xi_{\nu''})}^{b_{\nu''}} F(x)\,dx \right| < \epsilon. \tag{14}$$

The series (4) converges, and there is an integer ν_ϵ such that

$$\sum_{\nu>\nu_\epsilon} O_\nu < \frac{\epsilon}{4}. \tag{15}$$

* By slight modifications the proof of the sequel may be made to cover also the case of the finite set Ξ.

The ν_ϵ integrals

$$\int_{a_\nu(\Xi_\nu)}^{b_\nu} F(x)\,dx \qquad (\nu = 1, 2, \cdots, \nu_\epsilon)$$

exist, and there is a number δ_ϵ such that

$$(16) \qquad \delta_\epsilon < |a_\nu - b_\nu|, \quad \left|\left(\int_{a_\nu(\Xi_\nu)}^{b_\nu} - \int_{a'_\nu}^{b'_\nu}\right) F(x)\,dx\right| < \frac{\epsilon}{2\nu_\epsilon} \qquad (\nu = 1, 2, \cdots, \nu_\epsilon)$$

if throughout

$$(17) \qquad |a_\nu - b_\nu| - |a'_\nu - b'_\nu| < \delta_\epsilon.$$

This number δ_ϵ is the number desired: if I is any interval-set enclosing Ξ narrowly and of length $D_I < \delta_\epsilon$ the inequality (14) will hold. For

$$(18) \qquad D_I = \sum_{\nu'} \left(|a_{\nu'} - b_{\nu'}| - |a'_{\nu'} - b'_{\nu'}|\right) + \sum_{\nu''} |a_{\nu''} - b_{\nu''}|,$$

and so

$$(19) \qquad |a_{\nu'} - b_{\nu'}| - |a'_{\nu'} - b'_{\nu'}| < \delta_\epsilon, \quad |a_{\nu''} - b_{\nu''}| < \delta_\epsilon.$$

By comparison of (16) and (19) we see that the indices ν' include the first ν_ϵ indices ν, $\nu = 1, 2, \cdots, \nu_\epsilon$. Denoting the left member of (14) by L, we have

$$L \leqq \sum_{\nu=1}^{\nu_\epsilon} \left|\left(\int_{a_\nu(\Xi_\nu)}^{b_\nu} - \int_{a'_\nu}^{b'_\nu}\right) F(x)\,dx\right| + \sum_{\nu > \nu_\epsilon} \left|\int_{a_\nu(\Xi_\nu)}^{b_\nu} F(x)\,dx\right|$$

$$+ \sum_{\nu' > \nu_\epsilon} \left|\int_{a'_{\nu'}}^{b'_{\nu'}} F(x)\,dx\right|,$$

so that by (15) and (16) in view of (19) we have indeed the desired inequality (14)

$$L < \epsilon.$$

3°. The preceding theorem reduces the problem of construction of all narrow Ξ-integrals essentially to the corresponding problem for the case $\Xi = (b)$, and this remark holds likewise if the integrand function $F(x)$ is required to be on ab continuous except at points $x = \xi$.

In 5° I exhibit by a known process an essentially narrow or non-absolutely convergent integral

$$(20) \qquad \int_{a\,((b))}^{b} F(x)\,dx,$$

and in connection with it develop examples designed to show the error of certain statements referred to in the introduction.

In 4° I use such an integral (20), for the case $ab = 01$, as the element for a simple construction according to the conditions of 2° of a Ξ-integral for the general set Ξ of ab, non-absolutely convergent in the neighborhood of every point ξ.

4°. We denote by $F(x) = \psi_{01}(x)$ a function on the x-interval 01 continuous except at $x = 1$ for which the integral

$$\int_{0\,((1))}^{1} F(x)\,dx$$

converges non-absolutely. A function $F(x) = \psi_{ab}(x)$ of similar character on the interval ab $(a \gtrless b)$ is given by the transformation

$$\psi_{ab}(x') = \psi_{01}(x), \qquad x' = (b - a)x + a.$$

Denoting the middle point of ab by c, we denote by $\chi_{ab}(x)$ the function on ab everywhere continuous except at a and b which on ac is $\psi_{ca}(x)$ and on cb is $\psi_{cb}(x)$. The functions $\chi_{ab}(x)$ are derivable from $\chi_{01}(x)$ by the transformation used above. The integral

$$\int_{a\,((a,\,b))}^{b} \chi_{ab}(x)\,dx$$

converges, non-absolutely near a and b.

Then with respect to any set Ξ (1°) we construct a function $F(x)$ everywhere continuous on ab except at its singularities ξ and such that the integral

$$\int_{a\,(\Xi)}^{b} F(x)\,dx$$

converges non-absolutely near every point ξ. The function $F(x)$ is on $a_\nu b_\nu$ the function $\chi_{a_\nu b_\nu}(x)$, or ψ_{ab_ν} if $a_\nu = a =$ not-ξ, or ψ_{ba_ν} if $b_\nu = b =$ not-ξ.

The effectiveness of this construction follows easily from the theorem of 2° since the oscillations O_{ab}, O_{01}, O'_{ab}, of the integrals

$$\int_{a\,((b))}^{X} \psi_{ab}(x)\,dx, \quad \int_{0\,((1))}^{X} \psi_{01}(x)\,dx, \quad \int_{a\,(a,\,b)}^{X} \chi_{ab}(x)\,dx$$

on the respective intervals ab, 01, ab have the relations

$$O_{ab} = O_{01}|a - b|, \quad O'_{ab} \leqq O_{ca} + O_{cb} = O_{ab}.$$

5°. *Construction of a function $F(x)$ everywhere continuous on ab except at b, for which the integral*

$$\int_{a\,((b))}^{b} F(x)\,dx \tag{20}$$

converges non-absolutely.

We introduce any non-absolutely convergent series

$$\sum_{\nu=1}^{\infty} u_\nu, \tag{21}$$

—for instance, the modified harmonic series $(u_\nu = -(-1)^\nu/\nu)$,—and further any infinite sequence H of intervals $h_\nu = a_\nu b_\nu (\nu = 1, 2, 3, \cdots)$ lying progressively on ab, two consecutive intervals having no common inner point, and b being the limit-point of each of the sequences $(a_1, a_2, \cdots)$, $(b_1, b_2, \cdots)$.

Then we determine the function $F(x)$ as follows. On the interval h_ν $F(x)$ is everywhere continuous, does not change sign, and vanishes at the extremities $a_\nu b_\nu$, and further

$$\int_{a_\nu}^{b_\nu} F(x)\,dx = u_\nu, \tag{22}$$

—for instance, $y = F(x)$ is given graphically by the two sides of the isosceles triangle of (signed) height $u_\nu/2(b_\nu - a_\nu)$ on the base $a_\nu b_\nu$. At points x of ab but of no interval $a_\nu b_\nu$ $F(x)$ vanishes.

Then indeed $F(x)$ is everywhere continuous on ab except at $x = b$. In view of (22) $|F(x)|$ has on h_ν an upper limit greater than $|u_\nu|/|a_\nu - b_\nu|$, and hence, since the series (21) is non-absolutely convergent, $F(x)$ has values indefinitely great positive and negative in every neighborhood of $x = b$.

The narrow (b)-integral (20) is by definition (§ 2)

$$\underset{X=b}{\mathrm{L}^{X \text{ on } ab}} \int_a^X F(x)\,dx. \tag{23}$$

Now we have

$$\int_a^X F(x)\,dx = \sum_{\nu=1}^{n} u_\nu + \theta\, u_{n+1}, \tag{24}$$

where* X lies on the interval $b_n b_{n+1}$ and θ is some number, $0 \leqq \theta \leqq 1$. Hence it follows easily that

$$\int_{a\ ((b))}^{b} F(x)\,dx = \sum_{\nu=1}^{\infty} u_\nu, \tag{25}$$

and similarly that $|F(x)|$ is not $((b))$-integrable from a to b, since the series $\sum_\nu |u_\nu|$ diverges.

Thus, the integral (20) is non-absolutely convergent, and so (§ 4 III) it is an essentially narrow (b)-integral.

---

* We set $b_0 = a$ and $\sum_{\nu=1}^{0} u_\nu = 0$.

6°. *Exhibition for the* (b)*-integral* (20) *with respect to any two positive numbers* $\epsilon\delta$ *of an interval-set* J *not containing the point* b, *for which*

$$D_J < \delta, \quad \left| \int_J F(x)\,dx \right| > \epsilon.— \tag{26}$$

An example contravening HARNACK'S *theorem* 2.

In accordance with the footnote to note 1 of the first definition of § 2 this exhibition furnishes a more direct proof that the (b)-integral (20) is essentially narrow. And it evidently serves to disprove HARNACK'S theorem 2.

In view of the convergence for $\nu = \infty$ of the intervals h_ν to b there is an integer ν_0 such that

$$\sum_{\nu > \nu_0} |a_\nu - b_\nu| < \delta. \tag{27}$$

The series

$$\sum_{\nu > \nu_0} u_\nu$$

converges non-absolutely, and so contains a finite number (m) of positive terms $u_{\nu'} (\nu' = \nu_l > \nu_0;\ l = 1, 2, \cdots, m)$ such that

$$\sum_{\nu} u_{\nu'} > \epsilon. \tag{28}$$

For the interval-set J consisting of the m intervals $h_{\nu'}$ the inequalities (26) follow from (27) and (28).

7°. *Exhibition for the* (Ξ)*-integral* (20) $\left(\Xi = (b)\right)$ *of a set* H *closed and of content zero such that the corresponding* $(\Xi + \mathrm{H})$*-integral does not exist.*

The set H of all points $a_\nu b_\nu$ $(\nu = 1, 2, 3, \cdots)$ and the point b has the desired property. This follows directly from the exhibition of 6° in view of § 3 XIII′.

This example serves to show the presence of errors in certain statements of HARNACK* and BRODÉN* concerning the possibility of modifying the functional values of the integrand function at points of a set of content zero without altering the determination of the integral.

8°. *An example contravening* JORDAN'S *theorem* 3.

We have by 7° for the function $F(x)$ of (20) the (Ξ)-integral existent and the $(\Xi + \mathrm{H})$-integral non-existent, and furthermore $\Xi + \mathrm{H} = \mathrm{H}$.

Now we take by 4° a function $G(x)$ for which the $(\mathrm{H}) = (\Xi + \mathrm{H})$-integral exists. Then the $(\Xi + \mathrm{H})$-integral of $F(x) + G(x)$ is non-existent; for its existence would by § 3 II and XVII imply the existence of the $(\Xi + \mathrm{H})$-integral of $F(x)$.

9°. *An example contravening* HARNACK'S *theorem* 3.

The example of 8° is effective. For, if we set

---

*Mathematische Annalen, vol. 24 (1884), p. 220, and vol. 52 (1899), p. 222.

$$J(x) = \int_{a\,(\Xi)}^{x} F(x)\,dx, \quad K(x) = \int_{a\,(\mathrm{H})}^{x} G(x)\,dx, \tag{29}$$

$$L(x) = J(x) + K(x), \tag{30}$$

the functions $J(x)$, $K(x)$, $L(x)$ are everywhere continuous on ab; further

$$L(x'') - L(x') = \int_{x'}^{x''} \big(F(x) + G(x)\big)\,dx$$

for every interval $x'x''$ of ab containing no singularity ζ of $F(x) + G(x)$; and the set of singularities $\mathrm{Z} = \Xi + \mathrm{H}$ is reducible; and yet by 8° the HARNACK integral

$$\int_{a\,(\mathrm{Z})}^{b} \big(F(x) + G(x)\big)\,dx \tag{31}$$

does not exist.

The HÖLDER integral (31) however exists and its value is $L(b)$.

The non-existent integral

$$\int_{a\,(\Xi+\mathrm{H})}^{b} F(x)\,dx \tag{32}$$

would afford a simpler example from the point of view of this paper. However from the original HARNACK point of view one would object that the set of singularities of $F(x)$ is Ξ and not $\Xi + \mathrm{H}$.

In the example given the integrand function $F(x) + G(x)$ is everywhere continuous on ab except at its singularities ζ.

THE UNIVERSITY OF CHICAGO,
June 12, 1901.

ZUR LINEAREN TRANSFORMATION DER ϑ-REIHEN*

VON

F. MERTENS

§ 1.

Es sei ω eine complexe Grösse, in welcher der Coefficient von i positiv ist, und zur Abkürzung

$$q = e^{i\pi\omega}.$$

Definirt man die ϑ-Reihen durch die Gleichungen:

$$\vartheta_0(x, \omega) = \sum (-1)^n q^{n^2} e^{2n\pi ix},$$

$$\vartheta_1(x, \omega) = -i \sum (-1)^n q^{(n+\frac{1}{2})^2} e^{(2n+1)\pi ix},$$

$$\vartheta_2(x, \omega) = \sum q^{(n+\frac{1}{2})^2} e^{(2n+1)\pi ix},$$

$$\vartheta_3(x, \omega) = \sum q^{n^2} e^{2n\pi ix},$$

$$(n = 0, \pm 1, \pm 2, \pm 3, \cdots)$$

so besteht die lineare Transformation dieser Reihen in der Ersetzung von ω durch $(\gamma + \delta\omega)/(\alpha + \beta\omega)$, wo $\alpha, \beta, \gamma, \delta$ ganze der Gleichung

$$\alpha\delta - \beta\gamma = 1$$

genügende Zahlen sind.

Wird zur Abkürzung

$$\alpha + \beta\omega = N, \quad \gamma + \delta\omega = M$$

gesetzt, so kann mit Hilfe der elementaren Transformationsformeln:

$$\vartheta_0(x, 1 + \omega) = \vartheta_3(x, \omega),$$

$$\vartheta_1(x, 1 + \omega) = e^{\frac{i\pi}{4}} \vartheta_1(x, \omega),$$

$$\vartheta_2(x, 1 + \omega) = e^{\frac{i\pi}{4}} \vartheta_2(x, \omega),$$

$$\vartheta_3(x, 1 + \omega) = \vartheta_0(x, \omega),$$

$$\vartheta_0\left(\frac{x}{\omega}, -\frac{1}{\omega}\right) = \sqrt{-i\omega}\, e^{\frac{i\pi x^2}{\omega}} \vartheta_2(x, \omega),$$

* Presented to the Society (Chicago) April 6, 1901. Received for publication March 5, 1901.

$$\vartheta_1\left(\frac{x}{\omega}, -\frac{1}{\omega}\right) = -i\sqrt{-i\omega}\, e^{\frac{i\pi x^2}{\omega}}\, \vartheta_1(x, \omega),$$

$$\vartheta_2\left(\frac{x}{\omega}, -\frac{1}{\omega}\right) = \sqrt{-i\omega}\, e^{\frac{i\pi x^2}{\omega}}\, \vartheta_0(x, \omega),$$

$$\vartheta_3\left(\frac{x}{\omega}, -\frac{1}{\omega}\right) = \sqrt{-i\omega}\, e^{\frac{i\pi x^2}{\omega}}\, \vartheta_3(x, \omega),$$

leicht gezeigt werden, dass $\vartheta_1(x/N, M/N)$ die Gestalt $a e^{bx^2} \vartheta_1(x, \omega)$ hat, wo a, b Constanten bezeichnen.

Ist nämlich β nicht $= 0$, so sei β' der kleinste nicht negative Rest von δ in Bezug auf den Theiler β und

$$\delta = g\beta + \beta', \qquad \gamma = ga + a',$$

$$\gamma' = -a, \qquad \delta' = -\beta,$$

$$a' + \beta'\omega = N', \qquad \gamma' + \delta'\omega = M'.$$

Es ist dann

$$a'\delta' - \beta'\gamma' = 1, \quad |\beta'| < |\beta|,$$

$$\vartheta_1\left(\frac{x}{N}, \frac{M}{N}\right) = \vartheta_1\left(\frac{-x}{M'}, g - \frac{N'}{M'}\right) = -e^{\frac{g\pi i}{4}}\, \vartheta_1\left(\frac{x}{M'}, \frac{-N'}{M'}\right)$$

$$= ie^{\frac{g\pi i}{4}} \sqrt{\frac{-iM'}{N'}}\, e^{\frac{i\pi x^2}{M'N'}}\, \vartheta_1\left(\frac{x}{N'}, \frac{M'}{N'}\right);$$

$\vartheta_1(x/N, M/N)$ wird also die behauptete Gestalt haben, wenn $\vartheta_1(x/N', M'/N')$ sie hat. Die Ausdrücke M', N' sind aber insofern einfacher wie M, N, als $|\beta'| < |\beta|$ ist.

Ist β' noch nicht $= 0$, so können ebenso Ausdrücke

$$M'' = \gamma'' + \delta''\omega, \qquad N'' = a'' + \beta''\omega$$

angegeben werden, in welchen $|\beta''| < |\beta'|$ ist, und $\vartheta_1(x/N, M/N)$ hat zugleich mit $\vartheta_1(x/N'', M''/N'')$ die behauptete Gestalt.

Durch Fortsetzung dieser Schlüsse gelangt man zu einer linearen Transformation $(\gamma^{(m)} + \delta^{(m)}\omega)/(a^{(m)} + \beta^{(m)}\omega)$, in welcher $\beta^{(m)} = 0$ ist. Es genügt daher die behauptete Gestalt von $\vartheta_1(x/N, M/N)$ für den Fall $\beta = 0$ nachzuweisen.

Der Gleichung

$$a\delta = 1$$

zufolge ist dann $a = \delta = \pm 1$ und daher

$$\vartheta_1\left(\frac{x}{N}, \frac{M}{N}\right) = \vartheta_1(ax, a\gamma + \omega) = a\, e^{\frac{a\gamma\pi i}{4}}\, \vartheta_1(x, \omega).$$

Um die Constante b zu bestimmen, ersetze man x durch $x + N$. Es wird

$$\vartheta_1\left(\frac{x+N}{N}, \frac{M}{N}\right) = -\vartheta_1\left(\frac{x}{N}, \frac{M}{N}\right) = a\, e^{b(x+N)^2}\vartheta_1(x+N, \omega)$$

$$= a\, e^{bx^2+2bNx+bN^2}(-1)^{\alpha+\beta}\frac{\vartheta_1(x, \omega)}{q^{\beta^2}e^{2\beta\pi ix}}$$

und nach Forthebung von $a\, e^{bx^2}\vartheta_1(x, \omega)$

$$-1 = e^{2(bN-i\beta\pi)x}e^{bN^2}q^{-\beta^2}(-1)^{\alpha+\beta}.$$

Es muss also

$$bN - \beta\pi i = 0, \quad b = \frac{\beta\pi i}{N}$$

sein.

Die grösste Schwierigkeit bereitet die Bestimmung der Constante a. Differentiirt man die Gleichung

$$\vartheta_1\left(\frac{x}{N}, \frac{M}{N}\right) = a\, e^{bx^2}\vartheta_1(x, \omega)$$

nach x und setzt hierauf $x = 0$, so ergiebt sich

$$\frac{1}{N}\vartheta_1'\left(0, \frac{M}{N}\right) = a\,\vartheta_1'(0, \omega).$$

Es ist aber, wenn mit Herrn DEDEKIND

$$\eta(\omega) = q^{\frac{1}{12}}\Pi(1-q^{2n}) \qquad (n = 1, 2, 3, \cdots)$$

gesetzt wird,

$$\vartheta_1'(0, \omega) = 2\pi\eta^3(\omega)$$

und daher

$$a = \frac{1}{N}\left(\frac{\eta(M/N)}{\eta(\omega)}\right)^3.$$

Die Bestimmung von a ist hierdurch auf die von $\eta(M/N)$ zurückgeführt.

In dem Folgenden soll der Ausdruck $\eta(M/N)$ mit Hilfe der Formeln

$$\eta(1+\omega) = e^{\frac{i\pi}{12}}\eta(\omega),$$

$$\eta\left(-\frac{1}{\omega}\right) = \sqrt{-i\omega}\,\eta(\omega)$$

ermittelt werden.*

---

* HERMITE, Liouville's Journal, Ser. II., T. III.

DEDEKIND, *Erläuterungen zum XXVII Fragment in Riemann's Werken.*

H. WEBER, *Elliptische Funktionen und algebraische Zahlen* und *Zur Theorie der elliptischen Funktionen*, Acta Mathematica, Bd. VI.

§ 2.

Hilfssatz. Es seien A, B von Null verschiedene complexe Grössen, deren reelle Bestandtheile nicht negativ und nicht beide $= 0$ sind. Zieht man die Quadratwurzeln $\surd A$, $\surd B$ so aus, dass ihre reellen Bestandtheile positiv sind, so hat auch das Product $\surd A \surd B$ einen positiven reellen Bestandtheil und ist daher derjenige Werth der Quadratwurzel $\surd AB$, welcher dieselbe Eigenschaft hat.

Setzt man

$$\surd A = a + ia', \quad \surd B = b + ib',$$

so ist

$$A = a^2 - a'^2 + 2iaa', \quad B = b^2 - b'^2 + 2ibb',$$

$$\surd A \surd B = ab - a'b' + i(ab' + ba'),$$

und nach der Annahme a, b positiv, eine der Grössen $a^2 - a'^2$, $b^2 - b'^2$ positiv, die andere nicht negativ. Ist etwa $a^2 - a'^2$ positiv, so sind $a + a'$, $a - a'$ positiv. Da ferner die Grössen $b + b'$, $b - b'$ nicht entgegengesetzte Vorzeichen haben und ihre Summe positiv ist, so ist eine derselben positiv, die andere nicht negativ. Daher ist auch eine der Grössen

$$(a + a')(b - b'), \quad (a - a')(b + b')$$

positiv, die andere nicht negativ und ihre Summe $2(ab - a'b')$ somit positiv.

§ 3.

Es sei

$$a + \beta\omega = N, \quad \gamma + \delta\omega = M,$$

wo a, β, γ, δ ganze, der Gleichung

$$a\delta - \beta\gamma = 1$$

genügende, Zahlen sind, und es handele sich um die Bestimmung des Ausdrucks $\eta(M/N)$.

Es sind zwei Fälle zu unterscheiden, je nachdem $a\delta$ oder $\beta\gamma$ ungerade ist.

Sind a, δ ungerade, so darf a positiv angenommen werden, da man eintretendenfalls die Vorzeichen aller vier Zahlen a, β, γ, δ umkehren könnte, ohne den Quotienten M/N zu ändern.

Es sei zunächst β gerade und vom Null verschieden.

Der absolut kleinste Rest δ_1 der ungeraden Zahl $\delta - \beta(\gamma + 1)$ in Bezug auf den Theiler 2β ist ungerade und liegt demzufolge, vom Vorzeichen abgesehen, unter β. Setzt man

$$\delta = g\beta + \delta_1, \quad \gamma = ga + \gamma_1,$$

so ist

$$a\delta_1 - \beta\gamma_1 = 1$$

und γ_1 den Congruenzen

$$g \equiv \gamma + 1 \pmod 2,$$

$$\gamma_1 \equiv \gamma - g \pmod 2$$

zufolge ungerade. Es sei ferner β_1 der absolut kleinste Rest von β in Bezug auf den Theiler $2\delta_1$ und

$$\beta = h\delta_1 + \beta_1, \quad \alpha = h\gamma_1 + \alpha_1,$$

so dass h, β_1 gerade, α_1 ungerade sind.

Bezeichnet $\mathfrak{a}$ die Einheit mit dem Verzeichen von α_1 und setzt man

$$\alpha_1 = \mathfrak{a}\alpha', \quad \beta_1 = \mathfrak{a}\beta', \quad \gamma_1 = \mathfrak{a}\gamma', \quad \delta_1 = \mathfrak{a}\delta',$$

$$\alpha' + \beta'\omega = N', \quad \gamma' + \delta'\omega' = M',$$

so ist

$$\alpha'\delta' - \beta'\gamma' = 1,$$

α' positiv und die Bestimmung von $\eta(M/N)$ auf die von $\eta(M'/N')$ zurückführbar.

Man hat, wenn zur Abkürzung

$$\alpha_1 + \beta_1\omega = N_1, \quad \gamma_1 + \delta_1\omega = M_1,$$

$$e^{12} = \rho$$

gesetzt wird,

$$\eta\left(\frac{M}{N}\right) = \eta\left(g + \frac{M_1}{N}\right) = \rho^g \eta\left(\frac{M_1}{N}\right),$$

$$\eta\left(\frac{M_1}{N}\right) = \sqrt{\frac{iN}{M_1}}\,\eta\left(-\frac{N}{M_1}\right),$$

$$\eta\left(\frac{-N}{M_1}\right) = \eta\left(-h - \frac{N_1}{M_1}\right) = \rho^{-h}\eta\left(\frac{-N_1}{M_1}\right),$$

$$\eta\left(\frac{-N_1}{M_1}\right) = \sqrt{\frac{-iM_1}{N_1}}\,\eta\left(\frac{M_1}{N_1}\right) = \sqrt{\frac{-iM_1}{N_1}}\,\eta\left(\frac{M'}{N'}\right),$$

wo die Quadratwurzeln, wie überall in dem Folgenden, mit positivem reellen Bestandtheil auszuziehen sind. Hiernach wird

$$\eta\left(\frac{M}{N}\right) = \rho^{g-h}\sqrt{\frac{iN}{M_1}}\sqrt{\frac{-iM_1}{N_1}}\,\eta\left(\frac{M'}{N'}\right).$$

Bezeichnen $\mathfrak{c}$, $\mathfrak{d}$ die Einheit mit dem Vorzeichen von γ_1 beziehungsweise δ_1, so hat β der Gleichung $\beta = (\alpha\delta_1 - 1)/\gamma_1$ zufolge das Vorzeichen von $\mathfrak{cd}$ und jede

der Grössen $iN/M_1, -\mathfrak{i}\mathfrak{d}M_1, -i\mathfrak{c}\mathfrak{d}N$ einen positiven reellen Bestandtheil, daher ist

$$\sqrt{\frac{iN}{M_1}}\sqrt{-\mathfrak{i}\mathfrak{d}M_1} = \sqrt{\mathfrak{d}N},$$

$$\sqrt{-i\mathfrak{c}\mathfrak{d}N}\sqrt{\mathfrak{c}i} = \sqrt{\mathfrak{d}N}, \quad \sqrt{-i\mathfrak{c}\mathfrak{d}N}\sqrt{\mathfrak{c}\mathfrak{d}i} = \sqrt{N}$$

und sonach

$$\sqrt{\frac{iN}{M_1}} = \frac{\sqrt{N}}{\sqrt{-\mathfrak{i}\mathfrak{d}M_1}}\,\frac{\sqrt{\mathfrak{c}i}}{\sqrt{\mathfrak{c}\mathfrak{d}i}}.$$

Ist β_1 von Null verschieden, so hat es, der Gleichung $\beta_1 = (\alpha_1\delta_1 - 1)/\gamma_1$ zufolge, das Vorzeichen von $\alpha_1\delta_1/\gamma_1$ oder $\mathfrak{a}\mathfrak{c}\mathfrak{d}$ und die reellen Bestandtheile der Grössen $-iM_1/N_1, -1/\mathfrak{i}\mathfrak{d}M_1, -i\mathfrak{c}\mathfrak{d}N'$ sind positiv. Sonach ist

$$\sqrt{\frac{-iM_1}{N_1}}\sqrt{\frac{1}{-\mathfrak{i}\mathfrak{d}M_1}} = \frac{1}{\sqrt{\mathfrak{d}N_1}} = \frac{1}{\sqrt{\mathfrak{a}\mathfrak{d}N'}},$$

$$\sqrt{-i\mathfrak{c}\mathfrak{d}N'}\sqrt{\mathfrak{a}\mathfrak{c}i} = \sqrt{\mathfrak{a}\mathfrak{d}N'}, \quad \sqrt{-i\mathfrak{c}\mathfrak{d}N'}\sqrt{\mathfrak{c}\mathfrak{d}i} = \sqrt{N'},$$

und demzufolge

$$\sqrt{\frac{-iM_1}{N_1}} = \frac{\sqrt{-\mathfrak{i}\mathfrak{d}M_1}}{\sqrt{N'}}\,\frac{\sqrt{\mathfrak{c}\mathfrak{d}i}}{\sqrt{\mathfrak{a}\mathfrak{c}i}}.$$

Diese Gleichung ist aber auch noch richtig, wann $\beta_1 = 0$ ist, weil dann $\alpha_1\delta_1 = 1$, also

$$\alpha_1 = \delta_1 = \mathfrak{a} = \mathfrak{d}, \quad N' = 1,$$

$$\sqrt{\frac{-iM_1}{N_1}} = \sqrt{-\mathfrak{i}\mathfrak{d}M_1}, \quad \sqrt{\mathfrak{i}\mathfrak{a}\mathfrak{c}} = \sqrt{\mathfrak{c}\mathfrak{d}i}$$

ist.

Die Einsetzung der gefundenen Werthe von $\sqrt{iN/M_1}$, $\sqrt{-iM_1/N_1}$ ergiebt

$$\eta\left(\frac{M}{N}\right) = \frac{\rho^{g-h}\sqrt{\mathfrak{c}i}\sqrt{N}}{\sqrt{\mathfrak{a}\mathfrak{c}i}\sqrt{N'}}\,\eta\left(\frac{M'}{N'}\right);$$

da aber

$$\mathfrak{c}i = i^{\mathfrak{c}}, \quad \mathfrak{a}\mathfrak{c}i = i^{\mathfrak{a}\mathfrak{c}}$$

ist, so wird

$$\sqrt{\mathfrak{c}i} = e^{\frac{\mathfrak{c}\pi i}{4}}, \quad \sqrt{\mathfrak{a}\mathfrak{c}i} = e^{\frac{\mathfrak{a}\mathfrak{c}\pi i}{4}},$$

$$\frac{\sqrt{\mathfrak{c}i}}{\sqrt{\mathfrak{a}\mathfrak{c}i}} = e^{\frac{i\pi}{4}\mathfrak{c}(1-\mathfrak{a})} = i^{\frac{\mathfrak{c}(1-\mathfrak{a})}{2}}$$

und demnach

$$\frac{1}{\sqrt{N}}\,\eta\left(\frac{M}{N}\right) = \rho^{g-h}\,i^{\frac{\mathfrak{c}(1-\mathfrak{a})}{2}}\,\frac{1}{\sqrt{N'}}\,\eta\left(\frac{M'}{N'}\right).$$

Die Potenz

$$i^{\frac{c(1-a)}{2}}$$

lässt sich mit Hilfe des LEGENDRE-JACOBI'schen Zeichens vereinfachen.

Die Identität

$$\begin{aligned} c(1-a) &= (1-a)a'\gamma_1 + (a-1)(a'\gamma_1 - c) \\ &= (a'-a)\gamma_1 + (a-a_1)\gamma_1 + c(a-1)(a'c\gamma_1 - 1) \end{aligned}$$

ergiebt nach Ersetzung von $a - a_1$ durch $h\gamma_1$

$$c(1-a) = h\gamma_1^2 + (a'-a)\gamma_1 + c(a-1)(a'c\gamma_1 - 1),$$

und es wird daher, weil γ_1 ungerade ist,

$$c(1-a) \equiv h + (a'-a)\gamma_1 + (a-1)(a'c\gamma_1 - 1) \pmod 8,$$

$$i^{\frac{c(1-a)}{2}} = i^{\frac{h+(a'-a)\gamma_1}{2}} (-1)^{\frac{a-1}{2}\frac{a'c\gamma_1-1}{2}}.$$

Es ist aber

$$(-1)^{\frac{a-1}{2}\frac{a'c\gamma_1-1}{2}} = \left(\frac{a}{a'c\gamma_1}\right) = \left(\frac{a}{a'}\right)\left(\frac{a}{\gamma_1}\right),$$

$$i^{\frac{(a'-a)\gamma_1}{2}} = i^{\frac{a-a'}{2}} (-1)^{\frac{a'-1}{2}\frac{\gamma_1+1}{2}} (-1)^{\frac{a-1}{2}\frac{\gamma_1+1}{2}}$$

und nach dem quadratischen Reciprocitätssatze,

$$(-1)^{\frac{a-1}{2}\frac{\gamma_1+1}{2}} = \left(\frac{a'}{-\gamma_1}\right)\left(\frac{-\gamma_1}{a'}\right) = \left(\frac{a'}{\gamma_1}\right)\left(\frac{-\gamma_1}{a'}\right),$$

$$(-1)^{\frac{a-1}{2}\frac{\gamma_1+1}{2}} = \left(\frac{a}{-\gamma_1}\right)\left(\frac{-\gamma_1}{a}\right) \quad \left(\frac{a}{\gamma_1}\right)\left(\frac{-\gamma_1}{a}\right).$$

Aus den Gleichungen

$$a = h\gamma_1 + a_1,$$

$$-\beta\gamma_1 = 1 - a\delta_1, \quad -\beta_1\gamma_1 = 1 - a_1\delta_1$$

folgt überdies

$$\left(\frac{a}{\gamma_1}\right) = \left(\frac{a_1}{\gamma_1}\right) = \left(\frac{a\,a'}{\gamma_1}\right) = \left(\frac{a}{\gamma_1}\right)\left(\frac{a'}{\gamma_1}\right),$$

$$\left(\frac{-\beta\gamma_1}{a}\right) = \left(\frac{-\gamma_1}{a}\right)\left(\frac{\beta}{a}\right) = 1, \quad \left(\frac{-\gamma_1}{a}\right) = \left(\frac{\beta}{a}\right),$$

$$\left(\frac{-\beta_1\gamma_1}{a'}\right) = \left(\frac{-\gamma_1}{a'}\right)\left(\frac{\beta_1}{a'}\right) = 1, \quad \left(\frac{-\gamma_1}{a'}\right) = \left(\frac{\beta_1}{a'}\right) = \left(\frac{a}{a'}\right)\left(\frac{\beta'}{a'}\right).$$

Somit wird

$$i^{\frac{\varepsilon(1-\alpha)}{2}} = i^{\frac{h+a-a'}{2}} \left(\frac{\beta}{a}\right)\left(\frac{\beta'}{a'}\right).$$

Nach Ersetzung von $i^{h/2}$ durch ρ^{3h} ergiebt sich daher

$$\left(\frac{\beta}{a}\right)\frac{i^{\frac{1-a}{2}}}{\sqrt{N}}\ \eta\left(\frac{M}{N}\right) = \rho^{g+2h}\left(\frac{\beta'}{a'}\right)\frac{i^{\frac{1-a'}{2}}}{\sqrt{N'}}\ \eta\left(\frac{M'}{N'}\right).$$

Der Factor ρ^{g+2h} kann in zwei andere zerlegt werden, welche beziehungsweise nur von a, β, γ, δ und $a', \beta', \gamma', \delta'$ abhängen.

Der einfachste Weg zu dieser Zerlegung beruht auf dem Umstande, dass die Produkte $(a^2-1)(\beta^2-1), (\gamma_1^2-1)(\delta_1^2-1)$ Vielfache von 24 sind. Denn einer der Factoren a^2-1, β^2-1 muss nach dem FERMAT'schen Satze ein Vielfaches von 3 sein, weil a, β nicht beide durch 3 theilbar sein können, und a^2-1 ist ein Vielfaches von 8. Dieselben Gründe gelten für $(\gamma_1^2-1)(\delta_1^2-1)$. Es ist daher

$$\begin{aligned} 1 &\equiv a^2+\beta^2-a^2\beta^2 \\ 1 &\equiv \gamma_1^2+\delta_1^2-\gamma_1^2\delta_1^2 \end{aligned} \pmod{24},$$

und die Multiplication mit g, h ergiebt

$$\begin{aligned} g &\equiv a\,ga+\beta\,g\beta-a^2\beta\,g\beta \\ h &\equiv \gamma_1 h\gamma_1+\delta_1 h\delta_1-h\gamma_1^2\delta_1^2 \end{aligned} \pmod{24}.$$

Nach Ersetzung von $ga, g\beta, h\gamma_1, h\delta_1$, durch $\gamma-\gamma_1, \delta-\delta_1, a-a_1, \beta-\beta_1$ folgt hieraus

$$\begin{aligned} g &\equiv a(\gamma-\gamma_1)+\beta(\delta-\delta_1)-a^2\beta(\delta-\delta_1) \\ h &\equiv \gamma_1(a-a_1)+\delta_1(\beta-\beta_1)-h\gamma_1^2\delta_1^2 \end{aligned} \pmod{24},$$

und man erhält durch Addition

$$\begin{aligned} g+h \equiv a\gamma-a_1\gamma_1+\beta\delta-\beta_1\delta_1-a^2\beta\delta+a_1^2\beta_1\delta_1 \\ +\delta_1(a^2\beta-a_1^2\beta_1)-h\gamma_1^2\delta_1^2 \pmod{24}. \end{aligned}$$

Ersetzt man aber in den Identitäten

$$\begin{aligned} a^2\beta-a_1^2\beta_1 &= (a^2+aa_1+a_1^2)(\beta-\beta_1)+(a+a_1)(a\beta_1-a_1\beta), \\ a\beta-a_1\beta_1 &= (a+a_1)(\beta-\beta_1)+a\beta_1-a_1\beta_1 \end{aligned}$$

$\beta-\beta_1$ durch $h\delta_1$, $a\beta_1-a_1\beta$ durch $-h$, so wird

$$\begin{aligned} \delta_1(a^2\beta-a_1^2\beta_1) &= h\delta_1^2(a^2+aa_1+a_1^2)-h\delta_1(a+a_1), \\ a\beta-a_1\beta_1 &= h\delta_1(a+a_1)-h, \end{aligned}$$

und die Addition ergiebt

$$\delta_1(a^2\beta - a_1^2\beta_1) = -a\beta + a_1\beta_1 - h + h\delta_1^2(a^2 + aa_1 + a_1^2).$$

Mit Hilfe dieser Gleichung nimmt die vorstehende Congruenz für $g + h$, wenn zur Abkürzung

$$A = a\gamma + \beta\delta - a^2\beta\delta - a\beta,$$

$$A' = a'\gamma' + \beta'\delta' - a'^2\beta'\delta' - a'\beta',$$

$$C = h(a^2 + aa_1 + a_1^2 - \gamma_1^2)$$

gesetzt wird, die Gestalt

$$g + 2h \equiv A - A' + \delta_1^2 C \pmod{24}$$

an.

Nach Ersetzung von a durch $h\gamma_1 + a_1$ wird ferner

$$C = (h^3 - h)\gamma_1^2 + 3h^2a_1\gamma_1 + 3ha_1^2.$$

Es ist also einerseits, weil $h^3 - h$ nach dem FERMAT'schen Satze ein Vielfaches von 3 ist,

$$C \equiv 0 \pmod 3$$

und anderseits, weil h gerade, a_1 und γ_1 ungerade sind,

$$C \equiv -h + h^2 + 3h = 4\frac{h}{2}\left(\frac{h}{2} + 1\right) \pmod 8$$

$$\equiv 0 \pmod 8.$$

Da hiernach C ein Vielfaches von 24 ist, so folgt

$$g + 2h \equiv A - A' \pmod{24},$$

$$\rho^{g+2h} = \rho^{A-A'},$$

und man hat

$$\left(\frac{\beta}{a}\right)\frac{i^{\frac{1-\alpha}{2}}\rho^{-A}}{\sqrt{N}}\,\eta\left(\frac{M}{N}\right) = \left(\frac{\beta'}{a'}\right)\frac{i^{\frac{1-\alpha'}{2}}\rho^{-A'}}{\sqrt{N'}}\,\eta\left(\frac{M'}{N'}\right).$$

Diese Gleichung zeigt, dass der Ausdruck

$$\left(\frac{\beta}{a}\right)\frac{i^{\frac{1-\alpha}{2}}\rho^{-A}}{\sqrt{N}}\,\eta\left(\frac{M}{N}\right)$$

ungeändert bleibt, wenn man von den Zahlen a, β, γ, δ zu $a', \beta', \gamma', \delta'$ übergeht. Letztere Zahlen besitzen aber dieselben Eigenschaften, welche bei den

ursprünglichen Zahlen a, β, γ, δ vorausgesetzt werden und sind insofern einfacher, als $|\beta'| < |\beta|$ ist. Denn es ist

$$a'\delta' - \beta'\gamma' = 1,$$

a', δ' ungerade, β' gerade, a' positiv und $|\beta'| < |\delta_1| < |\beta|$.

Ist β' noch nicht $= 0$, so kann eine neue Reihe von vier Zahlen $a'', \beta'', \gamma'', \delta''$ ermittelt werden, für welche der vorstehende Ausdruck denselben Werth hat wie für a, β, γ, δ. Da die Zahlen $|\beta|, |\beta'|, \cdots$ eine abnehmende Reihe bilden, so führt dieses Verfahren nach einigen Schritten zu vier Zahlen $a^{(n)}, \beta^{(n)}, \gamma^{(n)}, \delta^{(n)}$, deren zweite $= 0$ ist.

Hiermit ist die Aufgabe in allen Fällen auf die Bestimmung des vorstehenden Ausdruckes in dem Falle $\beta = 0$ zurückgeführt. Dann ist aber

$$a = \delta = 1, \quad A = \gamma, \quad \left(\frac{\beta}{a}\right) = \left(\frac{0}{1}\right) = 1,$$

und der fragliche Ausdruck wird

$$= \rho^{-\gamma}\, \eta(\gamma + \omega) = \eta(\omega).$$

Somit ist bei geradem β

$$\eta\left(\frac{M}{N}\right) = \left(\frac{\beta}{a}\right) i^{\frac{a-1}{2}} \rho^A \sqrt{N}\, \eta(\omega).$$

Ist β ungerade, so ist $\beta - a$ gerade, $\delta - \gamma$ ungerade,

$$a(\delta - \gamma) - (\beta - a)\gamma = 1$$

und daher nach dem vorhergehenden Falle

$$\eta\left(\frac{M}{N}\right) = \eta\left[\frac{\gamma + \dfrac{(\delta - \gamma)\omega}{1 + \omega}}{a + \dfrac{(\beta - a)\omega}{1 + \omega}}\right] = \left(\frac{\beta - a}{a}\right) i^{\frac{a-1}{2}} \rho^C \sqrt{\frac{N}{1 + \omega}}\, \eta\left(\frac{\omega}{1 + \omega}\right),$$

wo

$$C = a\gamma + (\beta - a)(\delta - \gamma)(1 - a^2) - a(\beta - a).$$

Es ist aber

$$\left(\frac{\beta - a}{a}\right) = \left(\frac{\beta}{a}\right),$$

$$\begin{aligned} C - A - 1 &= (1 - a^2)(a\gamma - a\delta - \beta\gamma - 1) \\ &= (a^3 - a)(2\delta - \gamma) \\ &\equiv 0 \pmod{24}, \end{aligned}$$

weil $a^3 - a$ durch 3 und 8 theilbar ist, und

$$\eta\left(\frac{\omega}{1+\omega}\right) = \sqrt{\frac{i(1+\omega)}{\omega}}\,\eta\left(-1-\frac{1}{\omega}\right)$$

$$= \rho^{-1}\sqrt{\frac{i(1+\omega)}{\omega}}\,\eta\left(-\frac{1}{\omega}\right)$$

$$= \rho^{-1}\sqrt{\frac{i(1+\omega)}{\omega}}\,\sqrt{-i\omega}\,\eta(\omega).$$

Da überdies die Grössen $iN\omega, -i\omega(1+\omega), -i\omega$ positive reelle Bestandtheile haben, so wird

$$\sqrt{\frac{iN}{\omega}}\sqrt{\frac{-i\omega}{1+\omega}} = \sqrt{\frac{N}{1+\omega}}, \quad \sqrt{\frac{iN}{\omega}}\sqrt{-i\omega} = \sqrt{N}$$

und demnach

$$\sqrt{\frac{N}{1+\omega}} = \frac{\sqrt{N}}{\sqrt{-i\omega}}\sqrt{\frac{-i\omega}{1+\omega}}.$$

Daher ist, wie bei geradem β,

$$\eta\left(\frac{M}{N}\right) = \left(\frac{\beta}{a}\right) i^{\frac{a-1}{2}} \rho^{A}\sqrt{N}\,\eta(\omega).$$

§ 4.

Es sei $\beta\gamma$ ungerade. Man darf β positiv annehmen.
Man hat

$$\eta\left(\frac{M}{N}\right) = \eta\left(\frac{\delta+\gamma\frac{1}{\omega}}{\beta+a\frac{1}{\omega}}\right)$$

und daher auf Grund des bereits erledigten Falles

$$\eta\left(\frac{M}{N}\right) = \left(\frac{-a}{\beta}\right) i^{\frac{\beta-1}{2}} \rho^{B}\sqrt{\frac{N}{\omega}}\,\eta\left(\frac{-1}{\omega}\right),$$

wo

$$B = \beta\delta + a\gamma - a\gamma\beta^2 + a\beta$$

ist. Es ist aber

$$\left(\frac{-a}{\beta}\right) = \left(\frac{-1}{\beta}\right)\left(\frac{a}{\beta}\right) = (-1)^{\frac{\beta-1}{2}}\left(\frac{a}{\beta}\right),$$

$$\eta\left(\frac{-1}{\omega}\right) = \sqrt{-i\omega}\,\eta(\omega),$$

$$\sqrt{-iN}\,\frac{1}{\sqrt{-i\omega}} = \sqrt{\frac{N}{\omega}}$$

und sonach

$$\eta\left(\frac{M}{N}\right) = \left(\frac{a}{\beta}\right) i^{\frac{1-\beta}{2}}\,\rho^{\beta}\sqrt{-iN}\,\eta(\omega).$$